THE UNIVERSITY OF ALBERTA

# ANAPHORA PRODUCTION IN CANTONESE NARRATION

BY

© SIK LEE CHEUNG

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

PSYCHOLINGUISTICS

THE DEPARTMENT OF LINGUISTICS

EDMONTON, ALBERTA

FALL, 1986

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-32576-3

I am a part of all that I have met;
Yet all experience is an arch wherethrough
Gleams that untraveled world whose margin fades
Forever and forever when I move.
How dull it is to pause, to make an end,
To rust unburnished, not to shine in use!
As though to breathe were life! Life piled on life
Were all too little, and of one to me
Little remains; but every hour is saved
From that eternal silence, something more,
A bringer of new things; and vile it were
For some three suns to store and hoard myself,
And this gray spirit yearning in desire
To follow knowledge like a sinking star,
Beyond the utmost bound of human thought.

Tennyson, *Ulysses*

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Sik Lee Cheung

TITLE OF THESIS: Anaphora Production in Cantonese Narration

DEGREE: Master of Science

YEAR THIS DEGREE GRANTED: 1986

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

Signature: Cheung Sik Lee

Permanent Address: F712 Telford Gardens
Kowloon Bay
Hong Kong

Date: 16 Sept., 1986

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled Anaphora Production in Cantonese Narration submitted by Sik Lee Cheung in partial fulfilment of the requirements for the degree of Master of Science in Psycholinguistics.

Lois M. Stanford

Supervisor

Jean-Jues Lin

Hyeni Trube

John J. Hogan

Date 3 September 1986

To

the Memory of My Father

## ABSTRACT

This thesis investigates the psychological processes underlying the alternation of definite full noun phrases (NP's), pronouns and zero anaphora in the on-line production of narratives in Cantonese. Two major approaches to the study of the processes concerned are examined. One approach, the accessibility model, claims that the speaker varies the explicitness of anaphora during an on-going discourse on the basis of the following factors: (1) the referential distance between two successive mentions of a referent; (2) the potentiality of interference from other referents in the immediately preceding discourse. Positive correlations between the two factors and the explicitness of anaphora are postulated.

Twenty narratives produced by Cantonese speakers in Hong Kong based on two children's stories were analysed with respect to the two factors. Improved versions of the measurements designed by Givon (1983, 1984) to assess the factors in question were applied to the data. The results of the measurements indicated that the pronouns overlapped substantially with the zero anaphora. Moreover, the more explicit forms of anaphora in most cases did not show an advantage over the less explicit forms in occurring at high referential distances when the effect of potential interference was controlled. These results are not predicted by the accessibility model.

The other approach, the structural model, claims that the alternation among anaphoric forms is a function of the structure of discourse. One specific claim is that NP's are used to reinstate a referent at the beginning of an episode regardless of how recently the referent was last mentioned. The correlation of NP's with the beginnings of episodes is attributed to attentional factors.

Two experiments were done to test the correlation between NP's and episode boundaries. One experiment involved artificial episode boundaries which were created by inserting sequences of blank pages in two picture books. The other experiment used natural episode boundaries defined by one hundred independent judges. Forty narratives were elicited in each experiment and neither experiment found much evidence for the correlation. Instead, the following general principle was strongly implied in the data: the speakers avoided zero anaphora after an episode boundary and would use pronouns if this option is available. The principle provides a plausible

explanation for the overlapping between pronouns and zero anaphora which could not be accounted for by the accessibility model. It is concluded that episode boundaries which characterize the disruption of the flow of discourse have an effect on the alternation of pronouns and zero anaphora in Cantonese narration.

## ACKNOWLEDGEMENTS

I feel much gratitude to Dr. Lois Stanford, my thesis supervisor, who guided and supported me graciously over a period of two years during which time this thesis was conceived and completed. Her remarkable patience and enthusiasm in my work have been extremely important to me. I benefited a lot from the numerous discussions she had with me and appreciate very much her reviewing of the earlier drafts and her invaluable comments which have been incorporated into the final version of the thesis.

My special thanks go to the other members of the committee - Drs. Kyril Holden, Jenn-Shann Lin and John Hogan - for their discussions with me, their insights and valuable criticisms. I am grateful to Dr. Hogan for his advice on the statistical analyses of the studies in the thesis.

I am indebted to many friends for their help and support during the preparation of the thesis: Eric Lee, Fanny Kong, Sam Li, Suk Kam Chan, Jan Huen Wong, Yiu Chi Hui, Alice Chow, Seui Fun Yiu and Cheuk Wing Lam recruited the subjects for the experiments; Ljubov Reitman provided continued encouragement throughout and proofread the earlier drafts; Kok Tien Kwee kindly acted as my computer consultant; Ying Wu Wei's moral support and inspiration have been crucial and extraordinary.

For their participation and interest in the studies, I thank my subjects, who came from the following schools in Hong Kong: Kiangsu-Chekiang College, Queen Elizabeth School, Ying Wah School, Pui Ying Middle School, Our Lady's College, Holy Family College, Heep Tung College and Lai Tze College.

My thanks also go to the members of the Department of Linguistics at the University of Alberta and the TESL Centre at Concordia University. Their friendship and help have made my study in Canada the most interesting and rewarding experience in my life. I am especially grateful to Drs. Bruce Derwing and Matthew Dryer of the University of Alberta, and Drs. Patsy Lightbown and Alex Sharma of Concordia University, whose teaching has influenced me a great deal.

I want to express my deepest gratitude to my mother and brothers for their love, moral and financial support.

My master's study was made possible by teaching and research assistantships from the Department of Linguistics, University of Alberta, for which I am thankful.

TABLE OF CONTENTS


CHAPTER PAGE

1. OVERVIEW ..... 1
   1.1 Introduction ..... 1
   1.2 The Study ..... 2
   1.3 Overview ..... 2
   1.4 Terminology ..... 4
2. THE ACCESSIBILITY MODEL ..... 5
   2.1 Introduction ..... 5
   2.2 Listener-orientation ..... 6
      2.2.1 Inner Speech ..... 6
      2.2.2 Speech Error ..... 7
      2.2.3 Ambiguous References in Adult and Child Speech ..... 7
      2.2.4 Summary ..... 8
   2.3 Time and Interference ..... 8
      2.3.1 Givon's Cross-language Studies ..... 9
      2.3.2 Clancy's Pear Stories ..... 10
      2.3.3 Givon's "Referential Distance" ..... 12
      2.3.4 Measuring the Effect of Time on Anaphoric Choice in Clancy (1980) ..... 14
      2.3.5 Unit of R.D. Measurement ..... 18
      2.3.6 Quantifying Potential Interference in the Pear Stories ..... 24
      2.3.7 Givon's "Potential Interference" ..... 27
   2.4 Summary ..... 30
3. STUDY ONE ..... 32
   3.1 Stimulus Materials ..... 32
   3.2 Subjects ..... 32
   3.3 Procedures ..... 33
   3.4 Transcript ..... 35
   3.4 Measurements ..... 38
      3.4.1 Referential Distance ..... 38

3.4.2 Potential Interference..... 39
3.4.3 An Example..... 40
3.5 Results..... 42
3.5.1 Frequency Distributions of Anaphoric Options in Cantonese Chinese..... 42
3.5.2 Referential Distance and Potential Interference..... 47
3.5.2.1 Human Referents..... 47
3.5.2.2 Nonhuman Referents..... 53
3.5.2.3 Zero Anaphora - Subject and Object..... 68
3.5.2.4 Pronouns - Subject, Object and Possessive..... 69
3.5.2.5 NP's - Subject, Object and Possessive..... 70
3.5.2.6 Summary..... 72
3.6 Discussion..... 74
4. THE STRUCTURAL MODEL..... 82
4.1 Introduction..... 82
4.2 Tomlin's (1984-5) Episode-Attention Model..... 82
5. STUDIES TWO AND THREE..... 88
5.1 Study Two..... 88
5.1.1 Introduction..... 88
5.1.2 Stimulus Materials..... 88
5.1.3 Subjects..... 90
5.1.4 Procedures..... 91
5.1.5 Transcript..... 91
5.1.6 Results..... 92
5.1.6.1 Overall Results..... 92
5.1.6.2 Story M, First Telling..... 93
5.1.6.3 Story M, Second Telling..... 94
5.1.6.4 Story F, First Telling..... 95
5.1.6.5 Story F, Second Telling..... 96
5.1.7 Discussion..... 98
5.2 Study Three..... 101
5.2.1 Introduction..... 101
5.2.2 Episode Boundary Judgement..... 102

5.2.2.1 Stimulus Materials..... 102
5.2.2.2 Subjects..... 102
5.2.2.3 Procedures..... 103
5.2.2.3 Results..... 103
5.2.3 Anaphor Search..... 104
5.2.4 Results and Conclusions..... 105
6. SUMMARY AND GENERAL DISCUSSION..... 109

REFERENCES..... 112
APPENDIX A: Stories M and F..... 116
APPENDIX B: Frequency Distributions of Human Pronouns with Quantifiers..... 123
APPENDIX C: A Representative Narrative Elicited with the Cartoon Videotape from Tomlin (1984-5)..... 124

## LIST OF TABLES


| TABLE | DESCRIPTION | PAGE |
|---|---|---|
| 1 | Age Distribution of Narrators - Study One | 32 |
| 2 | Frequency Distributions of Anaphoric Options in Spoken Cantonese | 43 |
| 3 | Test of Significance of Difference in the Proportons of Each Anaphoric Option Between the Frog and the Monster | 44 |
| 4 | Test of Significance of Difference in the Proportions of Each Anaphoric Type for Human and Nonhuman Referents | 45 |
| 5 | Zero Anaphora - Human | 47 |
| 6 | Pronouns - Human | 48 |
| 7 | Frequency Distribution of Zero Anaphara and Pronouns for Human Referents from Distance "5" to "11[+]" on the P.I. Level of "1" | 49 |
| 8 | Full NP's - Human | 51 |
| 9 | Zero Anaphora - Frog | 54 |
| 10 | Zero Anaphora - Monster | 54 |
| 11 | Pronouns - Frog | 55 |
| 12 | Pronouns - Monster | 56 |
| 13 | Full NP's - Frog | 59 |
| 14 | Full NP's - Monster | 59 |
| 15 | Full NP's - Frog (With P.I.=2) | 66 |
| 16 | Full NP's - Monster (With P.I.=2) | 66 |
| 17 | Zero Anaphora - Subject (All Referents) | 68 |
| 18 | Zero Anaphora - Object (All Referents) | 69 |
| 19 | Pronouns - Possessive (All Referents) | 69 |
| 20 | Pronouns - Subject (All Referents) | 70 |
| 21 | Pronouns - Object (All Referents) | 70 |
| 22 | Full NP's - Subject (All Referents) | 71 |
| 23 | Full NP's - Object (All Referents) | 71 |
| 24 | Full NP's - Possessive and Adverbial Phrase (All Referents) | 71 |
| 25 | Frequency Distributions of Subject and Object NP's from | |

Distance "5" to "11[+]" on the P.I. level of "3"..... 72
26 Summary of Major Significant Differences in Proportions in Study 1.....73
27a Age Distribution of Experimental Group in Study Two..... 90
27b Sex Distribution of Experimental Group in Study Two..... 91
27c Education Level of Experimental Group in Study Two.....91
28 Frequency Distributions of Explicit and Non-explicit Types of Anaphora..... 92
29 Proportions of NP's and Non-explicit Anaphora in the First Telling of Story M.....93
30 Frequency Distributions of Explicit and Non-explicit Anaphoric Options in the First Telling of Story M..... 93
31 Proportions of NP's and Non-explicit Anaphora in the Second Telling of Story M.....94
32 Frequency Distributions of NP's, Pronouns and Zero Anaphora in the Seond Telling of Story M..... 94
33 Proportions of Explicit and Non-explicit Anaphoric Types in the First Telling of Story F.....95
34 Frequency Distributions of NP's, Pronouns and Zero Anaphora in the First Telling of Story F.....96
35 Proportions of Explicit and Non-explicit Types of Anaphora in the Second Telling of Story F.....97
36 Frequency Distributions of NP's, Pronouns and Zero Anaphora in the Second Telling of Story F.....97
37 Age Distribution of Episode Judges..... 103
38 Education Level of Episode Judges..... 103
39 Frequency Distribution of Anaphora - Study Three.....106

## LIST OF FIGURES


| FIGURES | DESCRIPTIOPN | PAGE |
|---|---|---|
| 1 | Serial Position Curves for Verbatim Recall in Experiment I of Jarvella (1971) | 19 |
| 2 | Schematic Representation of Picture Arrangement in Control Condition - Study Two | 88 |
| 3 | Schematic Representation of Picture Arrangement in Experimental Condition - Study Two | 89 |
| 4 | Episode Boundary Judgements on Story M | 104 |
| 5 | Episode Boundary Judgements on Story F | 105 |

# CHAPTER ONE

## OVERVIEW

### 1.1 Introduction

In telling stories, narrators tend to talk about referents introduced earlier in the story, and they typically have, at their disposal, a variety of linguistic forms to reinstate these referents. English speakers, for example, could re-introduce a referent such as *a toad* with the pronoun "it", or with full noun phrases like "the toad", and perhaps even "the prince under a magic spell." This form of reference, which points back to some previously-mentioned referents, is commonly known as **anaphora** in discourse analysis and psycholinguistics (Halliday and Hasan, 1976). The production of anaphora in natural speech is the primary concern of this thesis.

The full noun phrase, referred to as NP hereafter, and the pronoun are two common types of anaphora in the world's languages; the zero/null anaphor is another popular type, an example of which is illustrated by the Ø in the sentence "The toad received the magic kiss and Ø turned into a handsome prince."

The availability of more than one type of anaphoric expression implies that the speaker is probably constantly making decisions about whether to re-introduce a referent with a more or a less explicit anaphoric form in on-line anaphora production. Such decisions are unlikely to be purely random. The non-accidental nature of anaphoric choice can be demonstrated by replacing all the underlined items with *the man* in the following speech sample taken from Chafe (1980, p.301, Speaker 1):

(1) (a) ... And uh--...and then he gets down out of the tree,
(b) and he ... dumps all his pears into the basket,
(c) and the basket's full,
(d) and one of the pears drops down to the floor,
(e) and he ... picks it up,
(f) and he takes her kerchief off,
(h) and Ø places it in the basket which is ... very full.

What does the new version with the NP substitutes sound like? It does not result in syntactic ill-formedness; however, it has an artificial flavour to it, particularly when the underlined items are in close proximity.

If we reject the simple-minded notion that anaphoric choices are made in a purely ad hoc fashion, a legitimate question that follows is: what is the basis of such choices? Could it be the distance between two successive mentions of a referent hinted at above: the shorter the distance is, the less explicit an anaphor needs to be?

A purely distance-based account does not seem to take us too far before it runs into a problem. Reconsideration of the speech sample in (1) reveals a very distinct contrast: the speaker, on one hand, reintroduces the referent *basket* with an NP in (1c) although the last occurrence of *basket* is just one single word away; on the other hand, he uses a pronoun to reinstate *the man* in (1e) in spite of two intervening clauses. How can this kind of asymmetry be explained? Is this a common phenomenon or is it just a purely idiosyncratic feature? Furthermore, are the decisions involving anaphoric choices oriented to the needs of the speaker or to those of the listener? Answers to these questions will provide important insights into the processes underlying anaphoric production.

## 1.2 The Study

The study reported in this thesis is an experimental investigation of the psychological processes underlying anaphora production in discourse. In order to make this complex endeavour manageable, the investigation was restricted to one genre of discourse, namely the ***oral narrative***. The spoken mode rather than the written one was studied, because the former seems to be more revealing than the latter about the psychological processes underlying language production in general. Writing appears to be more "prone" to highly cultivated skills and factors such as the conscious stylistic intents of the writer which are highly idiosyncratic, and which constitute a special "marked" form of language.

The experimental work was carried out on ***Cantonese Chinese*** in order to make a comparison with English studies in the same domain, with the hope that cross-linguistic studies can help differentiate universal principles from language-specific ones. Cantonese Chinese, moreover, is an SVO language and it had not been studied so far with respect to the issue of concern here.

## 1.3 Overview

The thesis is organized into six chapters. Chapter One consists of an introduction and an overview of the thesis. Chapter Two contains a critical review of the ***accessibility*** model - one of the two most popular models in the study of anaphora

production. This model is primarily concerned with the time gap (referential distance) between two successive mentions of a referent and the presence/absence of potential interfering referents in the preceding discourse. Measurements used in the literature are carefully examined with respect to their underlying assumptions and their power in assessing the two factors of time gap and potential interference.

The accessibility model is evaluated experimentally in Chapter Three: twenty oral narratives based on two stories and produced by native Cantonese speakers in Hong Kong aged between 17 and 20 were collected. Improved versions of the measurements designed by Givon (1983) were performed on the narratives with the main goal of finding out how explanatory are the two factors, referential distance and potential interference, which make up the accessibility model.

Chapter Four deals with the second of the two popular models - the ***structural*** model, which claims that the alternation among anaphoric forms is a function of the structure of discourse - its thematic organization and its hierarchical structure. This model assumes that discourse is hierarchically structured into various levels. The number and definition of these levels may differ from one researcher to another, yet the following four levels are generally accepted: story, episode/paragraph, event/sentence and clause. A positive correlation between NP and episode boundaries has been proposed on cognitive basis by some researchers working within this framework (van Dijk,1982; Tomlin, 1984-5; Chafe, 1976; *inter alia*).

Two experiments designed to test the correlation between NP selection and episode boundaries are reported in Chapter Five. One experiment involved artificially defined episode boundaries. Following Tomlin (1984-5), the boundaries were created by interrupting the attention of the narrator while he/she was telling the story. A total of forty narratives were collected under this condition, and the narrators' anaphoric choices at these artificial boundaries were examined and compared against forty control samples which were produced in a minimally interrupted condition.

The second experiment involved independent episode judgements on the same stories used in the first experiment. The pictures were arranged in the form of cartoon strips. One hundred Cantonese speakers, drawn from the same pool of subjects in Hong Kong but not participants in the other experiments, acted as episode judges: they read the strips and then decided which groups of pictures formed a unified section. Their judgements of episodes were then used as the basis for testing the postulated correlation of NP choice and episode boundary. The twenty natural narrative samples examined in Chapter Three were used again in this study. Episode

boundaries were first marked in the transcripts of these narratives, and then the frequencies of the various anaphoric forms were calculated.

The final chapter consists of a general discussion of anaphora production on the basis of the results obtained from the various studies done in this thesis.

## 1.4 Terminology

Two comments about terminology are in order before moving into the main body of the thesis.

First, the expression "refer(erence)" will sometimes be used loosely to relate a non-explicit anaphor to its antecedent NP or referent. Strictly speaking, it is not correct to say that a non-explicit anaphor such as a pronoun refers to its referent or antecedent NP since non-explicit anaphora by themselves do not have identifying power. Some linguists prefer to use the term "co-refer(ence)" when talking about non-explicit anaphora; for example, "he" co-refers with "John". Such fine distinction between "refer(ence)" and "co-refer(ence)" is not always made here, particularly when the distinction results in a clumsy expression.

Second, the term "referent" is assumed to mean the *concept* of a referent. Therefore, an expression such as "the accessibility of a referent" should be interpreted as "the accessibility of the concept/mental representation of a referent".

# CHAPTER TWO

# THE ACCESSIBILITY MODEL

## 2.1 Introduction

Two major approaches are taken in the study of anaphora production at the level of discourse. The first approach, which is the concern of this chapter, can be called the ***accessibility*** model. It is a listener-oriented model which considers primarily the following two factors: first, the amount of ***time*** elapsed between two successive mentions of the same referent, and second, the potentiality of ***interference*** from other referents in the discourse immediately preceding the referent of concern. Givon (1983,1984) is the most enthusiastic proponent of this model.

The fundamental assumption of this model is that the speaker is constantly making assessment of how accessible a referent is to the hearer in an on-going discourse: the more accessible a referent is, the less coding material is required for the hearer to identify the referent, and consequently the less elaboration the speaker needs to make on his/her task (Chafe, 1974, 1976; Givon, 1983, 1984).

The basis on which the speaker makes the accessibility judgement lies in the two parameters of ***time*** and ***interference*** mentioned above. The speaker is assumed to be sensitive to the correlations in (2).

> (2) (a) The wider the gap between present and preceding appearance of the referent/topic in the discourse register, the harder it would be for the *hearer* to identfy it.
>
> (b) The more competing topics/referents are present in the immediately preceding discourse environment of a topic/referent, the harder it would be for the hearer to identify it.
>
> - Givon, 1984, p.112

Underlying the correlations is the assumption that hearers are constrained by the limited capacity of their working memory (Peterson and Peterson, 1959; Murdock, 1961; Baddeley and Hitch, 1974); they can only process a restricted amount of information at a time, and the information is subject to decay and interference.

In brief, the accessibility model is a listener-oriented model, which assumes that the speaker empathizes with the hearer's needs in identifying references, and that the speaker *strategically* assesses these needs on the basis of time and interference. Moreover, the assessments manifest themselves in the language product as anaphora

with varying degrees of explicitness.

Given these general claims and assumptions, more specific questions about the model can be raised:

(1) What are the justifications for assuming that the speaker atunes his speech to the needs of the hearer?

(2) Given that each anaphor is a manifestation of the speaker's judgement of how accessible a referent is at a given moment in a discourse, and that the judgement is based on the parameters of time and interference, is it possible to separate these two factors and measure them objectively in spoken production data? If yes, how? Specifically: what is the unit for measuring the effect of time? Is the unit psychologically real? What constitutes a competing referent? In what way do the two parameters interact? To what extent can they predict the alternation among the various anaphoric forms available in a specific language?

These questions will be dealt with one by one in the rest of this chapter, and a critical review of this model will also be given. To the first question we now turn.

## 2.2 Listener-orientation

Both listener- and speaker-orientation have independently been suggested in the literature to underlie the processes of anaphora production (Brown, 1983; Clancy, 1980; Marslen-Wilson, Levy & Tyler, 1982; *inter alia*). A thorough discussion of the latter orientation will be postponed till Chapter 4. The goal of the present section is to find evidence for the postulation that anaphora production, and perhaps the broader phenomenon of language production, is a listener-oriented process.

The listener-based postulation seems intuitively correct when we consider directly our experience as a speaker, with inner speech, speech errors and children's speech.

### 2.2.1 Inner Speech

Inner speech is the kind of speech we produce when we talk to ourselves in the absence of a second party - the listener. Just imagine that you are now talking to yourself: do you produce highly elaborated or abbreviated speech? It is perhaps not fair to ask yourself this question, since you are now conscious of the question and that might bias your production. The best way to study inner speech, presumably, is to catch yourself talking to yourself or catch others in the midst of their inner speech.

Vygotsky (1962) experimented on this interesting phenomenon and observed that

inner speech is characterized by "simplification and predication"; speakers tend to "leave out the subject and all words connected with it." (p.145) That implies that inner speech is full of zero anaphora. Why is it possible for speakers to do this? The answer, simple and straight-forward, is that the speaker and the listener coincide in this situation: the speaker knows what he himself is thinking, and the thoughts are no secret to the listener, who at the same time is the speaker. Therefore, as Vygotsky puts it, in inner speech 'a practically wordless "communication" [between the speaker and the hearer] of even the most complicated thoughts is the rule.' (p.145)

2.2.2 Speech Errors

Another piece of evidence in support of the assumption that anaphora production is a listener-oriented process comes from the so-called "speech errors" in spontaneous speech. Speakers, in spontaneous speech, quite often repeat and correct themselves (Maclay and Osgood, 1959). When they do so, they sometimes replace a less explicit anaphor with an explicit one, such as replacing *he* with *John* ; or sometimes they repeat an argument with elaboration at the conclusion of a syntactically complete utterance. An example is taken from the oral English data reported in Givon (1983, p.349):

(3) ...he wound up with fifty-nine sections of land
for a ranch is what *he* still owns there now,
*ol' John Leehugh...*

The repetition with elaboration suggests that the speaker upon finishing an utterance reflects that the identity of a certain referent might not have been made clear enough to the listener and hence decides to repeat the referent in a less attenuated form. This kind of repetition is sometimes called *afterthought* reference (Hyman, 1975), or is more commonly known as right-dislocation in linguistics (Givon, 1983).

2.2.3 Ambiguous References in Adult and Child Speech

Another source of evidence that supports a listener-based postulation can be derived from the ambiguous references that we very occasionally encounter in our daily conversations with adults, and quite often come across in the speech of very young children. Studies indicate that young children's speech tends to be full of pronouns, sometimes even when the references are entirely ambiguous to the listener

(Piaget,1926).

Apparently, young children are not conscious of a speaker-listener contract implied in all conversations, perhaps with the exception of inner speech. The contract requires that the speaker and the listener cooperate in making the communication effective and successful. One way the speaker contributes to the conversation is by being "clear" about what he/she is referring to (Grice, 1967).

Adult speakers are not always good at abiding by the principle of being clear. Chafe (1974), in his observations on anaphora production, noted that adult speakers exhibit "a certain degree of egocentricism"; they occasionally lose sight of the listener and make an anaphoric choice from their own perspective. As a result they use an anophor that is more attenuated than is normally expected. Furthermore, from our experience as listeners, we are probably justified in saying that speakers vary with respect to the degree of egocentricism: some speakers are more egocentric than others, and are therefore less clear about their references.

### 2.2.4 Summary

To sum up, evidence from inner speech, speech errors and children's speech does seem to favour the assumption that anaphora production is a listener-oriented process: we have observed that speakers can be non-explicit about their references in the absence of a second party, or when they are not conscious of the needs of the hearer; moreover, speakers repair their speech, from less explicit to more explicit reference, and sometimes tag an "afterthought" NP to the end of a syntactically complete utterance, when they decide that their original choice is not good enough for the hearer to access the referent.

The accessibility model, up to this point, seems to be a sound one with respect to its orientation. Our next set of questions is directed to its claim concerning *time* and *interference* .

## 2.3 Time and Interference

The accessibility model claims that the two parameters of time and interference play the most crucial roles in anaphora production; on the basis of the two parameters, the speaker judges the accessibility of a referent and varies the explicitness of an anaphor accordingly (Chafe, 1974, 1976; Clancy, 1980; Givon, 1983, 1984).

Our general questions are: Can we separately measure the effects of time and interference on anaphoric choice in oral production data? If yes, how? Answers to

these questions have been suggested in Clancy (1980) and Givon (1983, 1984), which are the foci of the rest of this chapter. What follows first is a brief introduction to these works, and then a critical review of the quantitative methodologies used in the studies.

### 2.3.1 Givon's Cross-language Studies

Givon's 1983 work is a cross-language study, whose main goal is to investigate the correlation between referent/topic accessibility (continuity, predictability) and various referent coding devices such as pronouns and NP's. Written and spoken texts from eight different languages form the data base, and the spoken texts consist of monologues and conversations. The eight languages are English, Spanish, Ute, Japanese, Bibilical Hebrew, Amharic, Hausa, and Chamorro.

In his 1984 work, Givon is concerned with the same correlation, using a totally different type of data: spoken English data from three adult speakers learning English as a second language at the initial stage.

The accessibility of a referent is considered by the researcher to be primarily determined by the two factors of time and potential interference, which can be quantified separately in both spoken and written texts. Two measurements are accordingly developed and applied to the first and second language data described earlier. A scale of referent accessibility is derived from the results of the application of the measurements, and the part of the scale that is most relevant to this thesis is reproduced below.

(4) Scale of Referent Accessibility/Predictability

*Most Accessible/Predictable Referents*
zero anaphora
unstressed pronouns
stressed/independent pronouns
Full NP's
*Least Accessible/Predictable Referents*

On the basis of the above scale, Givon (1984) postulates the following language universal:

(5) Quantity Universal
"More continuous, predictable, non-disruptive topics [referents] will be marked by *less marking material;* while less continuous,

> unpredictable/surprising, or disruptive topics [referents] will be marked by *more marking material*." (Emphasis original)
>
> - Givon, 1984, p.126

A preliminary comment on the universal is that it does not consider the type of language which "manage[s] reference in terms of a thematic policy," according to Grimes (1978, p.viii). This type of language distinguishes thematically important (central) referents from the unimportant ones: the central character of a story is introduced in a specific way and maintained by a specific set of anaphoric forms. Mambia is an example of this kind of language. Perrin (1978) reports that the zero anaphora in this language are reserved exclusively for the central character of a story and used no matter how many other referents intervene between the successive mentions of the central referent. On the other hand, only NP's are used to refer to non-central referents even in the absence of ambiguity or intervening material. A potentially interesting area for descriptive linguists to explore would be to find out how large this class of languages is.

It is important to point out that because of the presence of this kind of language whose speakers do not make their anaphoric choice on the basis of the two factors of time and interference, our discussion of these factors is necessarily restricted to languages that do not correlate their anaphoric types with specific sets of referents.

### 2.3.2 Clancy's Pear Stories

Clancy (1980) is a comparative study, investigating how Japanese and English speakers introduce and maintain referents in oral narration. In the study, the subjects were first asked to see a film which has no language in it, and which is about a pear-picker and the theft of his pears by a boy. They were then required to narrate the story to an interviewer who had seen the film before the interview. A total of forty narratives were collected and analysed with respect to the referential/anaphoric choice made by the subjects of each language group concerned.

The researcher considers time and interference the most important but not only factors that influence a speaker's referential/anaphoric choice in both languages. The other factors will be discussed in relation to the structural model. As for the two factors of time and interference, Clancy, like Givon, assumes that they can be separately measured.

A preliminary criticism concerning the experimental design of this study is in

order: the using of an interviewer who had prior knowledge of the story before the experiment resulted in undesirable effects in the results. Clancy reported that some of her story-tellers used definite expressions such as NP's and pronouns to refer to a character introduced for the first time in the narrative, and then used an indefinite noun phrase to reinstate the same character. She attributed this unusual language behaviour to the narrators' shifting between the "real" world and the "story" world during the process of telling the story.

In the real world, the narrators were aware that their interviewer had seen the film; therefore they could treat the story as old/known information, and refer to the characters mentioned for the first time in the interview with definite expressions. On the other hand, in the "story" world, the narrators were conscious that there were rules to conform to in telling a story; for example, new characters should be introduced with indefinite noun phrases. The excerpt in (6) best illustrates what Clancy described as the shifting between the two worlds of "reality" and "story".

(6) . . . A man was picking pears in . . . what seemed to be his orchard,
. . . and-. . . came along first,
. . /someone/ came along first.
. . . Someone came along before the kid on the bicycle
but I don't remember
who it was.
. . . Then a kid came along on a bicycle,

- Clancy, 1980, p.148

In the excerpt, the narrator introduced *the kid* with a definite noun phrase as if it were old information, and then he referred to him with an indefinite noun phrase as if the character were being introduced for the first time.

Whether the above kind of unusual behaviour was caused by world shifting or not is not very important to us; what most concerns us is the extent to which Clancy's data were confounded by this very artificial factor. This confounding variable could have easily been eliminated by using "real" listeners, who had not seen the film nor heard the story before the experiment.

In general, Clancy's study is relevant to the present thesis in the following respects: first, the same genre of discourse is studied; second, similar methodologies are used in forming the data base; namely, collecting samples of narratives by different speakers about the same story; third, there can be a meaningful comparison between the Cantonese data and those of English and Japanese provided by Clancy. The

anaphoric system of the Cantonese language differs from those of the English and the Japanese languages in an extremely interesting way: when non-human referents in oral narratives are concerned, Cantonese is more like Japanese than English in having only the zero anaphor and the NP; however, if the referent is human, Cantonese is closer to English in having a third type of anaphora - the pronoun.

Having provided a general background to the three works by Clancy and Givon, we now scrutinize how each of them quantifies the two factors of time and interference.

2.3.3 Givon's "Referential Distance"

The dimension of time is measured in terms of clauses in Givon (1983). The justification is that the clause, which is a grammaticalized proposition, "is the basic information processing unit in human discourse."(p.7)

The measurement, known as "Referential Distance" (R.D.), assesses the distance from the last mention of a referent to its current mention. "Mention" can either be overt or expressed by a zero anaphor. The minimal R.D. value assigned by Givon is 1, in which case the referent last occurs in the immediately preceding clause, and the highest value is arbitrarily set at "20" on the grounds that one cannot cope with infinity.

The typical procedures in applying this measurement are: first, a R.D. value is assigned to each token of anaphor; second, the values of all the tokens belonging to the same anaphoric type are averaged and reported. The prototypical value for zero anaphora is found to be "1" in Givon's cross-language study. NP's usually have the highest R.D. values while pronouns tend to fall between these two extremes, with stressed/independent pronouns appearing in the range of "1.7" and "2" and unstressed pronouns occurring between "1" and "1.2".

If those prototypical values with their associated types of anaphora are lined up in an increasing order, the scale of topic accessibility in (4) will almost be replicated with the exception that the unstressed pronouns partially overlap with the zero anaphora. It appears that the dimension of time alone, expressed in terms of referential distance, is nearly adequate to account for topic accessibility.

That is indeed what Givon seems to be suggesting: in the three analyses that he himself did - Givon (1984), and the two studies on Ute and spoken English collected in his 1983 volume, he does not apply to his data his measurement for assessing

potential interference. This is probably the reason why his colleague Tomlin (1984-5) interprets his model as a distance/recency model, dealing primarily with how recent the last mention of a referent is.

As far as the prototypical values are concerned, there is a serious problem with the procedure of averaging R.D. values, for the averages could be misleading. An excellent example in support of the last statement can be found in Sun and Givon (1985), which studies the word order of Mandarin Chinese by examining written and spoken texts in the language.

In the study, various types of anaphora are identifed and subjected to Givon's measurement of referential distance. An average of 1.8 is obtained for the post-verbal independent pronouns in written text. The mean falls right in the prototypical range for independent pronouns obtained in Givon (1983)! In general, the value suggests that independent pronouns in written Mandarin tend to occur two clauses away from the current clause containing the same referent. This "suggestion" turns out to be wrong!

The paper contains a breakdown of all the tokens within each anaphoric type. It is shown in the breakdown that thirty-four out of the total thirty-seven tokens of the type of pronouns of concern here actually have a R.D. value of "1", and the higher average is caused by two "outlaws", one having a R.D. value of "11" and the other having a value of "20[+]". Further examples like this one are readily available in the same study.

This problem is apparent in the study by Sun and Givon because the authors of this paper report a breakdown of R.D. values for each anaphoric category. It is important to point out that seven out of the nine studies collected in Givon (1983) report averages ONLY, and the scale of referent accessibility in (4) and the Quantity Universal in (5) are based in a very large part on these averages.

The same study can be used to illustrate a problem of a different sort: all the proponents of the accessibility model report their results of R.D. and potential interference (P.I.) separately; that is, the results are not cross-referenced, and this way of reporting the results makes it very difficult for the reader to interpret them in a meaningful way, and to evaluate the model in general. Take for example the NP's in the study: most of the categories of NP's analyzed in the study have about twenty per cent of their tokens showing a R.D. value of "1". One naturally suspects that it was potential ambiguity that was causing these explicit references to appear at the shortest referential distance. However, it is impossible to evaluate this because the P.I. results

are not cross-referenced with the R.D. values.

Let us illustrate the above point with a specific category of NP's, say, the postverbal proper names from written text. It is reported that 20% of the tokens recorded for this category have a R.D. value of "1", and that around 5.3% of the total tokens occur in a potentially ambiguous environment. Given these figures, it is impossible to tell how many of these 5.3% of tokens have a R.D. value of "1". If the amount is low, one would start asking why quite a substantial proportion of this most specific type of NP appears immediately after the last mention of their referents, and factors other than R.D. and P.I. might be found to be important in anaphora production. At the least, cross-referencing the R.D. and P.I. results would help the reader to evaluate how explanatory these two factors are; that is, how much of the data can be accounted for by these two factors alone.

To sum up, Givon measures the dimension of time in terms of the number of *clauses* from the left of the present clause to where the same referent last occurs. Two comments have been made regarding the typical procedures used by the proponents of the accessibility model: first, reporting average R.D. values is inadequate and could be misleading because the means are sometimes biased by certain "non-conforming" tokens; second, the failure to cross-reference the P.I. and R.D. results makes it difficult for the reader to evaluate the model.

2.3.4 Measuring the Effect of Time on Anaphoric Choice in Clancy (1980)

In her study, Clancy measures the dimension of time in terms of the number of *clauses* and the number of *sentences* intervening between two successive occurrences of a referent. The employment of two linguistic units for measuring the time gap between two successive mentions is justified from both the speaker's and the listener's perspectives. According to the researcher, the narrator does not make reference to a certain character on the basis of temporal intervals, say, at intervals of ten seconds; rather, the references "occur within discourse units of varying sizes at different points within the entire narrative" (p.133).

While it is not hard to accept Clancy's statement about arbitrary temporal intervals, it is not clear as to exactly what she is claiming in the quotation. That quotation does not say more than the following: connected speech, such as the narrative, tends to be made up of linguistically definable units some of which are recognized as discourse units, and references (in fact everything said in the discourse)

occur within discourse units. This does not in itself justify why references should be measured in terms of linguistic units.

More importantly, Clancy does not make specific her discourse units. This is easy to understand because if we pursue the question further, we will soon find out that the clause and the sentence are not the only units in the study of discourse; we can then challenge her by asking why the unit of temporal measurement is not done in terms of other discourse units, say, the episode. We can conclude that Clancy's justification for using linguistic units to measure the time gap between two successive occurrences of a referent from the speaker's perspective is inadequate.

From the perspective of the hearer, Clancy argues that hearing a story "does not involve simply the passage of time but rather the continuous cognitive task of processing the incoming information"(p.132); moreover, quoting the works by Fodor, Bever & Garret (1974), Jarvella (1971) and Bernardo (1976) without elaborating on them, she says that the psychological importance of the clause and that of the sentence as processing units in speech comprehension have independently been established.

The statement that hearing a story does not *merely* involve "the passage of time" is intuitively correct; however, her statement about the psychological status of the clause and the sentence as processing units is questionable: it is controversial as to whether it is the clause or the sentence that is the crucial unit for processing linguistic information. Sentences in this context refer to complex sentences.

It is wise for Clancy to adopt both linguistic constituents when it is not absolutely clear which is more appropriate. Nevertheless, it is important to point out two things: first, it is difficult to define sentences in oral data; second, it is hard to interpret her results derived by using the sentence as the measuring unit because she does not tell her readers how her sentences are defined. Are they defined on the basis of syntactic completeness, or sentence-final intonation? The question sounds funny because most of us would expect the two to coincide, but Chafe (1980) tells us that this is not necessarily the case in natural speech: "empirically we find in spontaneous speech a very frequent mismatch [between sentence-final intonation and syntactic completeness]." (p.30)

Chafe defines sentences in oral discourse on the basis of intonation. Defined on such basis, a sentence in the pear film narratives can be as brief as one word such as "Sure", or as long as the whole story if the narrator places the *only* sentence-final intonation at the very end of the narrative. The wide range of length that an

intonational sentence can have probably implies, among other things, a large amount of individual differences within and between speakers; moreover, it is not unreasonable to suspect that in some languages intonational cues to mark sentence boundaries may not be as important or as well understood as those in English. Clancy makes the comment in (7) with regard to individual differences:

> (7) In Japanese the effects of sentence boundaries are rather difficult to evaluate, since speakers showed *striking individual differences* in sentence structure. Some speakers seemed to place sentence boundaries in much the same manner as English speakers; others used extremely long sentences, occasionally even recounting the entire narrative in from one to three long sentences. Thus for certain Japanese speakers the ordinary size of discourse units was reversed and sentences, which in English constitute a kind of intermediate unit between the case frame and the episode, incorporated several different episodes. (Emphasis added.)
>
> - Clancy, 1980, p.143

Though it is not absolutely clear if Clancy defines her sentences on the same basis as Chafe does, it is logical to assume that she does. The quotation above does not suggest that her sentences are syntactic ones. Moreover, the papers by Clancy and Chafe are collected in the same volume, forming two parts of a collective work on the study of the "Cognitive, Cultural, and Linguistic Aspects of Narrative Production". Furthermore, the fragments of narratives that Clancy quotes are punctuated in exactly the same way as those contained in the appendix of the volume. Note that the narratives in this collective study are transcribed according to units defined by Chafe.

If Clancy's sentences are actually based on intonation rather than syntactics, it is crucial to point out that no studies dealing with the psychological status of sentences in speech comprehension define sentences on that basis; sentences in these studies are all defined syntactically. This is to say that her sentences do not have the kind of psychological support that she is claiming. Furthermore, one suspects the reliability of her results which were derived on the basis of using the sentence as the temporal unit, because the length of a sentence can differ from speaker to speaker and from one moment to another within the same speaker.

Could Clancy have defined her sentences on syntactic basis? Before answering the question, let us repeat in (8) the excerpt of narrative quoted in Chapter One.

> (8) (a) ... And uh--...and he then gets down out of the tree,
> (b) and he ... dumps all his pears into the basket,

(c) and the basket's full,
(d) and one of the pears drops down to the floor,
(e) and he ... picks it up,
(f) and he takes her kerchief off,
(g) and he ... wipes it off,
(h) and Ø places it in the basket which is ... very full.

The excerpt shows that the narrator conjoined a long chain of clauses with the word "and" to form a sentence. Linking clauses like this is in fact charateristic of the narratives analysed by Clancy and connected speeches studied by others, for example, the study done by Sydney and Pawley reported in Chafe (1980). Again, the problem with this kind of long, multiclausal sentences is twofold: first, its psychological status in comprehension is not studied in the literature, and so Clancy cannot claim for its pychological importance; second, it is subject to individual differences because the number of clauses conjoined may differ within and between subjects.

Does the above discussion concerning the sentence imply that one should or could only use the clause as the unit for measuring the effect of time on anaphoric choice in spoken discourse? The answer to this question deserves a separate section (Section 2.3.5) if one is seriously concerned with finding a unit of measurement that has psychological import.

Let us now examine the results reported by Clancy in relation to the effect of "distance" on referential choice. Since the dimension of time is measured in terms of linguistic units by Clancy, it is more appropriate to use the term "distance" than "time", and the same comment applies to Givon. The general trend observed in both the English and the Japanese sets of data is that the higher the number of intervening clauses is, the more likely a nominal reference is used. The increase is gradual and the most striking thing to note is that nominal references do not reach the one hundred per cent level until there are at least twenty-one clauses intervening between two mentions of the same referent. This figure is amazingly high considering that the other alternative for the Japanese speakers is the zero anaphora; that is , nothing.

Clancy's report of the results is detailed and systematic. Nevertheless, she does not escape the second criticism we raised against the studies done by Givon and his colleagues: the results of the two dimensions concerned (R.D. and P.I.) are not cross-referenced. It is reasonable to suspect that longer distance might just mean a higher chance of having competing referents in the intervening clauses. Moreover, there are again about twenty per cent of nominal references appearing at the minimal referential distance, and since R.D. and P.I. are not cross-referenced, it is impossible

to evaluate how many of these references are actually triggered by the effect of potential ambiguity.

To sum up, the dimension of distance (time) is measured in terms of the clause and/or the sentence by Clancy, and by Givon and his colleagues. The problem of defining sentences in oral data has been discussed. The overall result found is that the number of intervening clauses correlates positively with the explicitness of the anaphor, but the generality of the result is undermined by two problems: (1) reporting and basing the conclusions on averages alone; (2) failing to cross-reference the results of R.D. and P.I., which makes it hard for one to meaningfully evaluate the model. Having examined how the parameter of distance (time) is dealt with in the studies of concern, we now turn our attention to the question concerning units of measuring R.D. left off earlier in this section.

### 2.3.5 Unit of R.D. Measurement

The question raised earlier was whether one should use the clause alone as the unit for measuring R.D. The tentative solution proposed here is that one should use the clause as the basic unit of measurement AND that one should treat as clauses bi-clausal sentences whose two components are syntactically (and semantically) highly dependent on each other.

Let us justify our proposal by first reviewing the research done by Jarvella (1971) quoted by Clancy in support of the psychological status of the sentence as a processing unit in comprehension. In this study, two experiments designed to measure the capacity of the listener's working memory were performed. The subjects listened to narratives which were interrupted occasionally, and wrote down as much as they remembered from just before the interruption. In one experiment (II) the subjects were provided with cue words taken from the tested material and in the other experiment (I) no cues were given.

The stimulus material consisted of identical sequences of words in one of two syntactic configurations, shown in (9) with the figures in parentheses indicating the number of words in the clause and the letter "S" representing "sentence":

(9) (a) $S_1$[Clause $_1$ (7)] $S_2$[Clause $_2$ (6) + Clause $_3$ (7)]
(b) $S_1$[Clause $_1$ (7) + Clause $_2$ (6)] $S_2$[Clause $_3$ (7)]

In configuration (a), $S_1$ is a simple sentence and $S_2$ is a complex subordinated one; in

configuration (b), the final sentence is a simple one while the first sentence is a complex one.

The diagram below shows partial results of experiment I, in which cue words were not provided. Notice the sharp dip after Position 7, which indicates the boundary between C2 and C3.

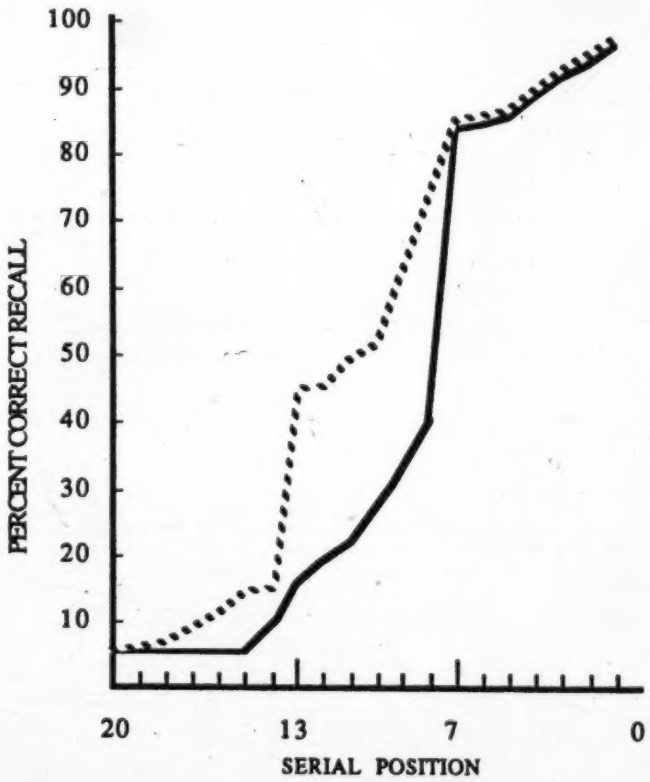



FIGURE 1: Serial Position Curves for Verbatim Recall in Exp I: Configuration A (doted line) and Configuration B (Solid line). (Adapted from Jarvella , 1971)

The sharp dip at Position 7 combined with a high percentage of correct recall at and above this position suggests that the last clause was stored as a unit in working memory and was best remembered. The last point has two important implications for the accessibility model as a listener-oriented model: first, the clause is a psychologically valid memory unit and hence an appropriate unit for measuring the effect of referential distance; second, the entities which occur in the final clause are readily available for current (co)-references. Chafe (1974, 1976), based on

introspective data, postulated that entities which are "in the addressee's consciousness" at a given moment in a discourse are most likely to be reinstated with attenuated anaphoric forms such as unstressed pronouns when potential interference is not at issue. The phrase "in the addressee's consiousness" makes sense only when it is taken to mean that the entities are in an active state; that is, they are in the addressee's working memory.

Similar results were replicated in Experiment II. In addition to the observations about the final clause, Jarvella observed the following point: the status of the second clause depended on whether it was part of $S_1$ or $S_2$: if it occurred in $S_2$, it would still be largely held in its surface syntactic structure, and be recalled *significantly* better than when it appeared in $S_1$. The researcher concluded on the basis of the last point that complex sentence boundaries are important and the sentence is a significant memory unit. In short, both the clause and the complex sentence were claimed to be critical units in language processing.

While Jarvella found evidence for the final complex sentence as an important processing unit in comprehension, other studies did not. The research done by Clark and Sengul (1979) belongs to the latter type of study.

Clark and Sengul (1979) reported two experiments which tested the availability of entities in the final clause and complex sentence for current (co)-reference. Their subjects were presented visually with context paragraphs, each consisting of three clauses in one of the two syntactic configurations schematized in (9) above. Instead of measuring verbatim memory, these two researchers measured the reading time of a target clause which was shown as soon as the subjects indicated that they had read and understood a paragraph. The target clause contained a pronoun or an NP whose (co)-reference had occurred in one of the three clauses of the paragraph. An actual test item is reproduced in (10) below.

(10) *Configuration* : $S_1[C_1]\ S_2[C_2+C_3]$

$C_1$: The campers erected a shelter of lightweight plastic in the clearing,

$C_2$: They piled a supply of dry wood next to a tree

$C_3$: while the smell of coffee mingled with smoke from the campfire.

Target: The tree was an enormous redwood.

In the above example, $C_2$ and $C_3$ form a complex sentence; in the other configuration ($S_1[C_1+C_2]$ $S_2[C_3]$), the first two clauses form a complex sentence instead, forming "Before the campers erected a shelter of lightweight plastic in the clearing, they piled a supply of dry wood next to a tree."

The results of both experiments show the following: the reading time of the target clause was *significantly* faster when the referent in question had been mentioned in $C_3$ rather than elsewhere; more importantly, when the referent had occurred in $C_2$, there was NO significant difference between whether the clause was part of $S_1$ or part of $S_2$; co-ocurring with the final clause in the same sentence did not give any privilege to the second clause in working memory.

On the basis of the above results and the assumption that reading time is a valid measure of comprehension speed, Clark and Sengul concluded that it is not the final sentence but the final clause which grants the entities it contains a privileged place in working memory; entities in the final clause are most readily available for current references. Moreover, the clause is a critical memory unit in speech comprehension while the complex sentence is not.

The study done by Jarvella and that done by Clark and Sengul came to conflicting results with respect to the status of the complex sentence as a crucial memory unit in speech processing. This could be due to differences in what they measured and differences in the experimental designs of the two studies. One of the differences appears to be of utmost importance to the controversy: in Jarvella's experiments, the two clauses that form a complex sentence are syntactically and semantically much more dependent on each other than those in Clark and Sengul's. Examples from the two studies are juxtaposed in (11) and (12) below.

(11) From Clark and Sengul (1979)

*Configuration (a):* $S_1[C_1]\ S_2[C_2+C_3]$

$C_1$: The campers erected a shelter of lightweight plastic in the clearing,
$C_2$: They piled a supply of dry wood next to a tree
$C_3$: while the smell of coffee mingled with smoke from the campfire.

*Configuration (b):* $S_1[C_1+C_2]\ S_2[C_3]$

$C_1$: Before the campers erected a shelter of lightweight plastic in the clearing,
C2: they piled a supply of dry wood next to a tree.
C3: The smell of coffee mingled with smoke from the campfire.

Target: The tree was an enormous redwood.

(12) From Jarvella (1971)

*Configuration (a):* $S_1[C_1]\ S_2[C_2+C_3]$

C1: The confidence of Kafach was not unfounded.
C2: To stack the meeting for McDonald,
C3: the union had even brought in outsiders.

*Configuration (b):* $S_1[C_1+C_2]\ S_2[C_3]$

C1: Kofach had been persuaded by the international
C2: to stack the meeting for McDonald.
C3: The union had even brought in outsiders.

Notice that the three clauses in Clark and Sengul's examples are all finite clauses and semantically independent of each other. In fact, they can be put in any order, and any two clauses can be subordinated without making much difference to the meaning of the paragraph; furthermore, $C_2$ is always a main clause. In contrast, Jarvella's clauses are more semantically related since they form a portion of a narrative; more importantly, the two clauses that form a complex sentence have a much closer syntactic and semantic bond than those of Clark and Sengul's, and in addition to the close bond, Jarvella's $C_2$ is always a subordinate clause.

The close bond between the two component clauses of Jarvella's complex sentence plus the subordinate status of $C_2$ implies that $C_2$ would be held in working memory longer in configuration (a), in which $C_2$ and $C_3$ form a sentence, than in

configuration (b); $C_2$ in (a) is very likely held until $C_3$ is heard so that a complete interpretation would be possible. The same principle does not apply to $C_2$ in (b) because the interpretation of $C_2$ in that case does not rely as much on $C_3$. This probably explains why Jarvella found a significant difference in verbatim recall of $C_2$ between the two configurations. Moreover, it is not hard to understand why no significant difference was found between the two configurations in Clark and Sengul: their C2 in both cases is quite independent syntactically and semantically.

Additional support for the phenomenon discussed above can be found in Jarvella and Pisoni (1970). In this study, within-sentence processing was studied and it was found that the order of the main and subordinate clauses was significant: subordinate-main order led to better immediate sentence recall than the reverse order did.

The general implication of the above discussion on the relationship between two clauses linked in a complex sentence is that the closer the semantic and syntactic bond is between the two clauses, the more appropriate it is to treat them as one single unit. The kind of bond demonstrated by Jarvella's complex sentences is not the closest because the verbatim recall of $C_2$ was much worse than $C_3$, even when $C_2$ was in the same sentence as $C_3$ (See Fig. 1). In English, one type of sentence that has a much closer bond between the component clauses is that containing sentential subject or object; for example: "Ron wanted to see a movie last night."

Last of all, two important comments are in order. First, though the final clause is generally recognized as the most accessible unit, it is important to recognize individual differences: Daneman and Carpenter (1980) did an excellent study which demonstrated quantitative differences in working memory between good and poor readers/listeners.

Second, a group of studies has demonstrated that the availability of an entity for (co)-reference is not determined by distance per se (Anderson, Garrod & Sanford, 1983; Frederiksen 1981; Purkiss, 1978; *inter alia*). In general, these studies found that the topical referent is more available than the non-topical ones to be referred to by a pronoun when both types of referents have been mentioned beyond the last clause.

Sanford and Garrod (1981) suggest that the representation of a topical referent remains longer in working memory than the representations of the non-topical ones. An obvious reason why this happens is that the topical referent is most expected to be talked about in subsequent discourse. The two researchers postulate that working

memory should contain at least the last clause and the representation of the topical referent, which would fade from memory only after sustained non-references in a relatively high number of subsequent clauses. It would be interesting for us to find out in the present thesis if the speaker is sensitive to the topical referent's special status and would use non-explicit anaphora to refer to topical referents even when their last mentions are relatively far away.

To sum up, two studies which looked into the roles of syntactic boundaries in working memory have been reviewed: both studies found that the clause is a valid memory unit and that the last clause is most accessible to the listener. We suggest that the clause be used as the basic unit for measuring R.D., and that sentences with syntactically and semantically highly dependent component clauses be treated as one unit. Other studies dealing with the same topic include Caplan (1972), Jarvella (1979) and Sach (1967, 1974). We now turn our attention to how potential interference is measured by Givon and Clancy.

### 2.3.6 Quantifying Potential Interference in the Pear Stories

In Clancy (1980), the dimension of potential interference is measured simply by tallying up the number of different competing referents that occur between two consecutive mentions of the same referent, disregarding the number of times each competing referent is mentioned. Moreover, a competing referent is defined by Clancy following Norman (1969) as one that is "maximally similar to" the referent of concern; that is, *similarity* is judged on the basis of factors such as "humanity" and "animacy".

To illustrate how the counting is carried out, let us take as an example the verb phrase "working in the pear tree": since the phrase can take any human referents who are capable of climbing up a tree and working there, all human referents that occur between this phrase and the last phrase containing the referent of concern should be included in the counting of competing referents. Notice that gender is not a consideration in Clancy's measurement.

Clancy's P.I. measurement is questionable in at least two respects. First, her definition of competing referents is over-conservative; it does not take into consideration other factors which play crucial roles in disambiguating references. Two extremely important factors are contextual information and world knowledge, which will be collapsed under the term "pragmatic information". Evidence for the importance of pragmatic information in resolving anaphora can readily be found in the real world.

One illustrative example reported in Marslen et al. (1982), which studies the oral narrative of one English speaker, is repeated in (13) below.

(13) "... so The Thing has to get down to ... the ground level ... before The Hulk recovers enough ... while *he* $_1$ still stunned ... so *he* $_2$ rips open the elevator doors and just sort of slides down the cable..."

(Marslen et al., 1982, p.363)

In the quotation above, there are two instances of the pronoun "he", and there are also two male creatures - The Thing and The Hulk, which can both take the pronoun "he", and which are both semantically compatible with the verb phrases headed by the two "he's". Are the two pronominal references ambiguous?

The answer is "no" because the excerpt occurs in the following context: the two referents were fighting on top of a skyscraper, and The Hulk has just fallen sixty storeys to the ground level below. Given this context, it must be The Hulk that is "stunned" after such a bad fall because our world knowledge tells us that the feeling of being "stunned" can be associated with falls; moreover, since the Hulk is lying stunned on the ground, it is unlikely that he at the same time can be ripping open the elevator doors. Therefore the second "he" must be co-referential with The Thing while the the first one co-refers with The Hulk.

The two pronominal references can in fact be resolved on the basis of world knowledge and the word "recovers" alone. The word suggests the idea of "in a bad state of health or mind", and "stunned" is a state of "not feeling well"; in other words, the word "recovers" alone can provide the right context for the interpretation of "who is stunned". Given this kind of contextual information, general knowledge about "recovers" and "stunned" AND the larger context described in the preceding paragraph, it is extremely unlikely that the listener could not resolve the two pronominal co-references.

The example in (13) illustrates that the listener can rely on various sources of information in the process of anaphora resolution at the level of discourse, and that the speaker is aware of these possibilities. It seems the speaker can manipulate the information density: he/she can provide just the right amount of information, or give the addressee more than is required for the identification of the intended referent. A very interesting question to ask is what tends to be the case in the real world: speakers being economical and providing the listener with just enough information, or being conservative and building in redundant information?

To sum up our first comment about Clancy's P.I. measurement, her definition of competing referents is extremely conservative; her relying entirely on the semantic compatibility between a referent and a verb phrase alone in determining if the referent is potentially interfering is under-estimating what the human processor can do.

The second problem with Clancy's P.I. measurement concerns her procedure of adding up the number of "competing" referents intervening between two successive mentions of a referent. Underlying this procedure is the assumption that the potentiality of interference is *directly* related to the number of intervening "competing" referents. The assumption seems to be supported by Clancy's study, which obtains a positive correlation between the number of intervening referents and the probability of using a nominal reference: the percentage of nominal references increases as the number of intervening referents increases until the latter hits the point of five, at which point the using of a nominal reference becomes obligatory. It seems it could be concluded from the result that the higher the number of intervening referents, the more likely it is to elicit a nominal reference. However, this conclusion is not warranted because there are frequent situations in which the number of intervening referents does not have much effect on anaphoric choice. Consider first the example in (14):

(14) John heard from his pen-pal Peter today. Peter has been working hard to promote communication among the members of the *International Pen-Pal Club*.

There is no intervening referent between the two mentions of the referent *Peter*, but an NP is strongly desirable to avoid potential ambiguity. One might argue that the stressed pronoun could be used to reinstate the second referent. Contrary to common belief, most listeners do not like non-explicit anaphora in a case in which subject switching is involved. Fletcher (1984) asked his subjects to listen to sentences such as the one in (15), with the blank filled with one of the four possibilities: an epithet such as "the guy", a stressed pronoun, an unstressed pronoun or a zero anaphor.

(15) Pete intended to go bowling with Sam last night but __ broke his leg.

By asking the question "Who broke his leg?", the researcher found that the stressed pronouns only had a twenty-eight per cent chance of being interpreted as indicating a change in subjects, and the probabilities for unstressed pronouns (19%) and zero anaphora (10%) were even lower.

Clancy indeed observed in her data that the switching of subjects "was a strong force in eliciting explicit forms of reference" (p.160). However, for some reason, this important phenomenon was not taken into consideration when she designed her measurement for potential interference.

We have observed that an NP can be strongly preferred in the absence of competing referents intervening two successve mentions of the same referent; however, on the other extreme, pronominal references are possible even if the number of intervening referents hits the sacred number of *five*. The example in (16) is one of such cases:

(16) Peter heard from his pen-pals *John, Paul, Tom, Ron* and *Don* today, and he was very pleased.

The pronominal co-reference in (16) is not ambiguous in spite of the presence of *five* intervening referents. One reason is that "Peter" is still the only possible antecedent for the pronoun "he"; the other male referents are somehow "locked in" to the group because of the appositive structure of their NP's. Another reason is that the listener can still remember the first referent by the time the pronoun is hit.

In short, our second comment about Clancy's P.I. measurement is that P.I. is not determined by the number of intervening referents *per se* . We have demonstrated that the discourse environment can be maximally ambiguous when there are no competing referents intervening between two successive mentions of a referent; on the other hand, the environment is not maximally ambiguous in spite of the presence of a relatively large number of intervening referents.

To sum up this section, Clancy's definition of intervening referents is inadequate, and her assessment of the potentiality of interference, which is done simply by tallying up the number of intervening "competing" referents, is highly questionable. We now examine how Givon and his colleagues quantify the parameter of potential interference.

### 2.3.7 Givon's "Potential Interference"

Potential interference is defined by Givon (1983) as "the disruptive effect which other referents within the immediately preceding register" have on the identification of the referent under concern. By "immediately preceding register", he refers to the distance *"between 1 and 5 clauses to the left"* of the clause concerned. P.I. is

measured by scanning to see if there are potentially interfering referents within the relevant distance: if no interfering referent is found, a value of "1" is assigned; if one or more competing referents are spotted, the value "2" is assigned.

Furthermore, a potentially interfering referent is defined as one that is syntactically and semantically compatible with the predicate/verb phrase of the clause under study. Some of the syntactic and semantic factors that are considered by Givon and his colleagues are animacy, agentivity, and plausibility as subject or object. Pragmatic information is not mentioned in Givon's introduction to his 1983 volume, but a few of his colleagues do comment that pragmatic information was taken into consideration in their analyses; for example Cooreman's study on Chamorro, and Bentivoglio's analysis of Spanish.

In contrast with Clancy's P.I. measurement, Givon's measurement does not assume that P.I. is directly related to the number of potentially interfering referents; his measurement is a binary function: either there is or there is not potential ambiguity. While his non-assumption is correct, his binary distinction is debatable, particularly when a language has more than one type of non-explicit anaphora.

Let us clarify the last point by first reconsidering the switch-subject case reported in (14), repeated in (17) below.

(17) John heard from his pen-pal Peter today. Peter has been working hard to promote communication among the members of the *International Pen-Pal Club*.

In the example, the second Peter is triggered by potential ambiguity. If we now replace Peter by a female referent, pronominal reinstatment becomes possible.

(18) John heard from his pen-pal Donna today. She has been working hard to promote communication among the members of the *International Pen-Pal Club*.

What has been illustrated by the two examples are two crucial points: first, according to Givon's P.I. measurement, both examples involve potential ambiguity, yet in one case an explicit anaphor is strongly preferable if not obligatory while in the other case a non-explicit one can readily be used; second, the two cases cannot be differentiated by R.D. since a R.D. value of 1 will be assigned to both anaphora - the second Peter and she. The two points combined imply that Givon's two measurements are not sensitive to the difference between (17) and (18).

One way to overcome this shortcoming in Givon's measurements is to modify his P.I. measurement when dealing with languages which have more than two anaphoric types: the measurement could be changed to a three- or four-level distinction depending on the specific properties of the language in question. For example, a value of "1" could be assigned to a syntactic slot when there is no potential ambiguity at all; that is, when the zero anaphor could be used to reinstate a referent in a particular clause without risking ambiguity; "3" is assigned to environments which are maximally ambiguous and in which the predicate concerned must take an NP in order to avoid ambiguity; and "2" is given to environments in a middle group: in these cases some of the features of the referents must be encoded to avoid ambiguity; for example, in English pronouns would fill the role by specifying number and gender.

Another difference between Givon's and Clancy's P.I. measurements is that Givon imposes a domain for searching referents - *"between 1 and 5 clauses to the left"* , most commonly 3 clauses to the left in the studies in his 1983 volume. He justifies this by suggesting that "there is a certain *decay effect* associated with referential distance of topics/referents"$_{(p.37)}$; that is, referents further away do not interfere as significantly with the identification of a current referent.

Givon's justification for imposing a search domain is correct when one is dealing with referents that have already occupied a position within the domain. Consider the pronoun in the example in (19):

> (19) My neighbours Mary and Diane took their dog to the park one evening. At the park, the dog went wild: it chased other dogs, fought with a cat, almost killed a bird and started chewing up plants. Finally, Diane said, "We should go home." She sounded very annoyed.

With its antecedent in the immediately preceding clause, the pronoun "she" is unlikely to be assumed to have the referent *Mary,* who appeared six clauses away. Givon's search domain seems to be justified in this case.

However, the search domain might run into a problem when the referent in question has not occupied a place in the domain, for example, the second mention of Diane. What is the P.I. value for it? It should be "1" according to Givon's measurement because all the referents in the search domain (preceding five clauses) do not interfere with Diane; those referents cannot speak (unless the dog is Pluto!). But, "1" is wrong in this case: there is potential interference from another human referent -

Mary, who occurs outside the domain.

The difficulty in assigning P.I. values to referents that have interfering referents outside the relevant distance suggests that we should not impose an arbitrary domain for searching potentially interfering referents. What an analyst could do is to take each token as an individual case and make a judgement on the basis of all kinds of information that are available - semantic, syntactic, morphological, thematic and pragmatic. A simple test would be to replace each token in turn with anaphora that are less explicit than the one actually used, and then to decide which is the optimal form for the identification of the referent in question.

## 2.4 Summary

The accessibility model has been reviewed with respect to the major studies in support of this model - Clancy (1980) and Givon (1983, 1984). The model is a listener-oriented one which postulates that the speaker constantly assesses the degree of accessibility of a referent to the listener. The degree of accessibility is inversely proportional to the amount of coding material required to mark the referent, with zero anaphora marking the most accessible referents, NP's coding the least accessible ones, and pronouns marking the referents in between the two extremes. Moreover, the degree of accessibility is determined mainly by the referential distance between two successive mentions of a referent and the potentiality of interference from other referents in the preceding discourse.

Measurements are proposed in the studies concerned to quantify the two factors of referential distance and potential interference. Problems with these measurements have been discussed and suggestions have been made on how to improve them. In measuring referential distance, it has been suggested that the basic unit of measurement be the clause (simple sentence) supplemented by complex sentences whose component clauses are highly semantically and syntactically dependent. In quantifying potential interference, a multi-level measurement has been suggested for languages that have more than two major types of anaphora, with lower levels designating lower potentiality of interference. It has been proposed that all sources of information be taken into consideration in determining the P.I. value for each token.

Furthermore, two points have been illustrated concerning how the results of the two measurements should be reported: first, the R.D. and P.I. values for all tokens should be cross-referenced so that one could judge if one or both factors are affecting a particular anaphoric choice; second, in reporting the R.D. values for the various types

of anaphora under study, the researcher should not merely report the averages for each class because the means could be misleading when there are atypical tokens present in the data.

The modified versions of the P.I. and R.D. measurements and the improved ways of reporting the results are applied to the data of our first study, which involves twenty narratives in spoken Cantonese. With improved methodologies, we hope that our study provides a better and more accurate evaluation of the accessibility model than what the previous studies have achieved. In our first study, we ask specifically the following question: "To what extent can the two factors of P.I. and R.D., which the accessibility model represents, account for anaphoric choice in the on-line production of narratives?"

# CHAPTER THREE

## STUDY ONE

### 3.1 Stimulus Materials

The stimulus materials consist of adaptions of two children's stories presented in pictures without words (Krahn, 1974; Mayer, 1974). One story (M) is about two boys who made a monster to fool various groups of villagers; and the other story (F) is about a frog's adventure at a fancy restaurant. Children stories were used because the pictures are self explanatory while adult cartoon strips are hard to interpret after removing the titles in the strips.

Each story consists of ten sets of pictures that are coloured and pasted in sequential order in a sketch book 37 cm x 29 cm in size. Each picture is about fourteen square centimeters in area, and two pictures are put side by side on one page. The two series of pictures for the stories can be found in a reduced size in Appendix A of the present thesis.

### 3.2 Subjects

Sixty volunteers were involved: twenty of them acted as the narrators while the rest of them were divided equally into two groups of story-listeners: the trial group and the experimental group. Among the story-tellers, nine were male and eleven were female. The age breakdown of the narrators is shown in Table 1.

| AGE | 17 | 18 | 19 | 20 |
|---|---|---|---|---|
| NUMBER | 6 | 9 | 4 | 1 |

Table 1: Age Distribution of Narrators.

The subjects came from three co-educational high schools in Hong Kong. Eighteen of the narrators and thirty-six of the listeners were in pre-university programs (equivalent to Grades 12 and 13 in North America) when the study took place in the

fall of 1985. The rest of the subjects were attending Form V at that time (i.e., Grade 11).

Information concerning the narrators' first and second languages/dialects was obtained. All the narrators speak Cantonese Chinese as their first language. As for other Chinese languages and dialects, three of the narrators claimed that they could speak fluent Mandarin, and one of these three speakers said that she used the language frequently at home; moreover, one speaker reported that she could speak a different dialect of the Yue language, to which Cantonese belongs. No non-Cantonese accent was detected in these subjects.

All the subjects have a considerable amount of knowledge in English, which is used as the medium of instruction in their schools and over ninety per cent of the high schools in Hong Kong. They used the language extensively in reading and in doing their school assignments, but they all reported that they did not speak the language much, particularly outside the classroom.

Questions were also asked of the narrators as to whether they told fictional stories in their daily lives. One narrator said that she occasionally did that to her peers while the rest of the tellers reported that they did not usually do so.

### 3.3 Procedures

Subjects were formed into groups of three: the grouping was done entirely on voluntary basis; what happened most often was that three friends or classmates put themselves into a group. The advantage of this way of grouping is that the subjects did not normally feel uneasy at the testing because they tended to be familiar with the other subjects in the same group.

Each group of subjects was interviewed separately by the experimenter before the testing and assigned randomly one of the two picture books; that is, each group only told one of the two stories. During the interview, the experimenter explained to the group that she was studying how people tell and remember stories. One of the subjects in the group would be required to be the story-teller whose task was to tell twice a story based on the picture book assigned to them; the other two subjects would take turns to listen to the story and recall the story immediately afterwards. Their tellings and recalls would all be recorded and studied.

Moreover, the subjects were told that the narrator would be following the picture book while telling the story, whereas the listeners would not be allowed to look at the

pictures. It was explained to them that the difference between the first and second tellings was that the story would unfold for the narrator during the first telling, but during the second telling, he/she would know the plot ahead of time.

After explaining the general procedures, the experimenter showed the subjects a sample picture book and asked if they had any questions about the procedures. When the procedures were clear to the subjects, a portable cassette recorder was turned on, and then the experimenter and the second listener retreated to a different room/location. The study usually took place in a quiet room or corner of the subjects' schools. Sometimes it was done in the experimenter's or a subject's home.

While the second listeners were waiting for their turns to listen to the story, the experimenter obtained from them as much background information as possible about themselves and their testing partners. Information that the second listeners could not provide concerning their partners was filled out at the end of the testing by asking the particular subjects concerned.

In most of the tellings, the narrator told the story to a listener of the same sex; in five cases in the first telling, the narrator told the story to a listener of the opposite sex, and the number was reduced to four in the second telling. No difference was observed between the same-sex and different-sex conditions.

Three points are in order with respect to the experimental design employed in this study. First, *real* listeners who had not heard the stories were used in both the first and the second tellings in order to create a true sense of communication, which is essential to the study of anaphora production. Recall the "world shifting" that Clancy (1980) observed of her narrators as a result of having a listener who had seen the film: some of the narrators introduced a referent in definite expressions when they were dominated by the awareness that the listener knew the story; yet when they remembered their role as a story-teller, they re-introduced the same character in indefinite noun phrases.

Second, there were a variety of purposes for recording the listeners' recalls: (1) to test their comprehension of the story; (2) to make sure that they paid attention to the story; (3) to make them feel that they had an active role to play in the experiment.

Third, three reasons were involved as to why the narrators were required to look at the picture books while telling the story. The first reason was to remove the memory constraint on the narrator. The next purpose was to have some, but not absolute, control over the narratives produced by the story-tellers. It was anticipated

that some of the subjects might produce radically different versions of the stories if they were not following the pictures because of potential individual differences in creativity and/or memory capabilities. The third purpose was to provide some direct test of the orientation of the accessibility model. By allowing the narrators to see the referents in the pictures, one could argue convincingly that anaphora production is a listener-oriented process if the narrators varied the explicitness of the anaphora as dictated by the needs of the listeners.

### 3.4 Transcript

Only the second tellings were used for this study because these tellings were obtained in contexts which most resembled the communicative contexts in real life: genuine listeners were involved and the narrators knew the story prior to the telling.

The second tellings were transcribed by the experimenter largely in standard Chinese characters. Since Cantonese is not normally used for writing, not all sounds in Cantonese have corresponding characters. When this kind of gap happened in the transcribing process, one of the following three devices was employed, depending on which one was most available to the transcriber in a particular case: (1) using phonetic symbols; (2) substituting for the character a homonyn; (3) making up characters on the basis of the general principles in forming Chinese characters (Wieger, 1964).

The most common way to form a Chinese character is to combine the meaning component and the sound component of a character. Take the character "to carry (things)" as an example: "carrying (things)" involves the using of the "hand(s)", and so the radical "hand" is appropriate for this character; as for the phonetic component, the character sounds most like that for "strength"; combining the radical for "hand" with that for "strength", we create a meaning-phonetic compound for "to carry (things)".

All linguistic information in the recordings was transcribed. Along with complete utterances about the story were included false starts, repairs, filled pauses, listeners' comments, and exchanges between the narrators and the listeners. Non-linguistic information such as "laughter" was noted in the transcript. Unfilled pauses that occurred when the narrator was turning the pages of the picture book were also measured.

To test for consistency in transcribing, the first five pages of a sample were transcribed again two weeks after the first transcription was done. The two versions

were compared and found to be very close to identical, the difference basically being the spacing of the words.

When all the samples had been transcribed, the transcripts were submitted to a *unit analysis*. To qualify as a unit, the utterance must belong to one of the following six classes of structures. The largest class is made up of clauses with a main verb; it typically contains simple sentences and conjoined clauses. One of the other five classes is verbless clauses with adjectival predicates; for example:

(19) *k euih-deih taai ngai la*
3-pl too short PRT
"They (are) too short."

In general, all Chinese languages do not require the stative verb "to be" between the subject and an adjectival predicate. Another category is complex sentences that contained a relative clause, or a sentential subject/object; the relative clause was considered part of an NP instead of an independent clause because its function is to identify referents. The next two classes consist of complex sentences in one of these two structures: (1) pivotal construction; (2) serial verb construction, an example of each is given below, with NP denoting the postition of an argument, and the schematic representations adapted from Li and Thompson (1981).

(20) Pivotal Construction
(a) *Schematic Representation*
(NP) $V_1$ NP $V_2$ (NP)
Direct Obj Subj

(b) *Example*
*ihng-dou ni-ek-gwaai-sau ne keih*
make this-CL-monster PRT stand
*haih go seui go-douh.*
at CL water there.

"(The boys) made the monster stand
in the water."

(21) Serial Verb Construction
(a) *Schematic Representation*
(NP) $V_1$ (NP) (NP) $V_2$ (NP)

(b) *Example*

*wan tiuh-sing bond-yuh go-seui-pok*
find CL-string tie-PRT CL-water-bubble

"(They) found a piece of string to tie the life-buoy."

CL in the gloss represents nominal classifiers, which precede most noun phrases in Cantonese. The characteristic of the pivotal construction is that the direct object of the first verb is also the subject of the second verb. In serial construction, series of (NP) V (NP) are simply juxtaposed.

The last class that was considered an independent unit was a negligible number of "Yes/No" answers to questions about the story: these phonologically brief units were considered single units by themselves because they express complete ideas - either confirmation or negation of a proposition.

When the units had been identified and numbered, the transcripts were then subjected to *an anaphor search*. Anaphora used to reinstate the following three (sets of) characters were singled out: (1) the two boys in Story M, who represent central human referents; (2) the monster in the same story, which is an inanimate referent; (3) the frog in Story F, which stands for a central, non-human, animate referent.

The anaphora were classified into three major classes and marked in distinctively different ways in the transcripts: the classes were zero anaphora, pronouns and NP's.

Zero anaphora were identified by examining if there were arguments missing in a unit. Two examples are given in (22) below to illustrate how zero anaphora were identified, with Ø marking explicitly the positions of the zero anaphora.

(22) (a) *Zero Subject*
*Ø kam-jyuh go-go-muhn*
cover-PRT DEM-CL-door.
"(They) covered that door."

(b) *Zero Object*
*ji-hau ne keuih-deih sau-faan Ø ji-hau ne, . . .*
then 3-pl take-back after PRT
"then, after they had taken back (the monster), . . . "

Two types of pronouns were considered - third person and possessive pronouns. The third person pronoun in spoken Cantonese is ***keuih*** "he/she", which is a legitimate pronoun for human referents regardless of sex, though it is found to be occasionally used with non-human referents. This could be attributed to two reasons: (1) influence from the written Chinese, in which a three-way gender distinction is

made on the character - the human male, the human female, and the nonhuman; (2) interference from English, particularly in Hong Kong, in which the language is one of the two official languages of the city.

The plural form of the pronoun is formed by attaching the plural suffix *deih* to the pronoun root, forming *keuih-deih* "they". The possessive pronoun is formed in two ways: one way is simply juxtaposing the pronoun root and the noun root, for example, *keuih bah-ba* " his father"; in the other way, a nominal classifier is inserted between the pronoun root and the noun root, for example, *keuih bun syu* " his CL book". There is a subtle difference between the two ways of forming the possessive pronoun: when the noun root is a non-human noun, it is obligatory to use the second way with the classifier; if the noun root is a kinship term, both ways are possible with a preference for the first way.

When the anaphor search had been completed, the tokens within each class were numbered. Tokens which occurred within "false start" and the speech followed by "repair" were not counted; for example, the underlined tokens in (23):

(23) (a) *False Start*
*(keuih-deih keuih-deih jeung) yan-waih hung-luhng-ga*
(3-pl 3pl OM) because dinosaur-frame
*hou gou*
very high

"(They they) Because the dinosaur frame is very high.

(b) *Repair*
*hau-meih keuih-deih jik haih go-euhng-go -sai-loh-jai*
afterwards 3-pl that is those-two-CL-children
*beih gau*
PASSIVE rescue

"In the end they$_j$ er, that is, those two children$_j$
were rescued."

After the tokens had been tallied, they were subjected to two measurements, which are elaborated in the following section.

## 3.4 Measurements

### 3.4.1 Referential Distance

Givon's (1983) R.D. measurement was applied to the tokens. This

measurement assesses the number of clauses to the left of the clause containing the referent concerned in which the last mention of a referent occurs. In this thesis, a minimal R.D. value of "0" was assigned if the two mentions occurred in the same clause. This happened when one of the mentions was in the relative clause, which was not counted as a separate unit, or in the possessive role.

3.4.2 Potential Interference

An improved version of Givon's (1983) P.I. measurement was performed on the data. It had three levels, each corresponding to one class of anaphora in Cantonese: Level 1 - zero anaphora; Level 2 - pronouns; Level 3 - NP's. The levels were assigned by asking the following question about each token: "On the basis of all sources of information available from the current and preceding clauses/units, which is the minimal level for the unique identification of the referent in question?" An example will be given to illustrate how the P.I. and R.D. measurements were applied to our data in the following section.

Two comments are in order regarding the assignment of P.I. levels: first, the zero anaphora and NP's for nonhuman referents could not be assigned the P.I. level of "2" because pronouns are not supposed to be an option for this type of referent; second, the P.I. level of "1" was not assigned to structures that do not allow deletion of arguments. There are two such structures: the pivotal construction and structures with coverbs.

The pivotal construction, which is illustrated in (20) and repeated in (24) below, does not generally allow ellipsis of the shared argument; that is, the "monster" in the example:

(24) *lihng-dou ni-ek-gwaai-sau ne keih*
make this-CL-monster PRT stand

*haih go-eui go-douh.*
at CL-water there.

"(The boys) made the monster stand
in the water."

The second type of structure that does not allow ellipsis of arguments consists of coverbs. Coverbs are functional words which originate from verbs. Some of them

still preserve their verbal quality; for example, the word *juhng* "use/with" can be used as a verb meaning "to use" and as a preposition analogous to the preposition "with" as in the English expression "He cut the apple with a knife."

Other coverbs have lost all their verbal qualities and become pure functional elements. One such example is the coverb *jeung*, whose counterpart in Mandarin Chinese has been classified as the object marker (OM) by some linguists (Li and Thompson, 1981).

There are good reasons for treating *jeung* as an object marker in Cantonese as well. First, the word, which comes from the verb "to marshall", is no longer an active verb in spoken Cantonese; the only time it can still be used as a verb is when a chess-player announces checkmate. Second, the function of this word in contemporary Cantonese is to highlight the direct object by fronting it to a preverbal position. An example is provided in (25) below:

(25) *jeung keuih-deih daai faan heui sehng-shi*
OM 3-pl bring back go city
"(The searchers) took them back to the city.

Notice that the main verb "bring" is preceded by the direct object "them" in the above example, which reverses the unmarked VO order in Cantonese.

*Jeung* is the most common coverb encountered in the present study. Other common ones are *hai* "from/at" as in "at a restaurant" and *tuhng(-maaih)* "with" as in "he went to the restaurant with his family."

### 3.4.3 An Example

An example is given in this section to illustrate how the R.D. and P.I. measurements were applied to the data. Consider first the following excerpt from the narrative produced by one of the story-tellers (Speaker 139):

(26) (a) *go-go gwai-fu sihk-gan choi ge sih-hauh ne,*
that lady eat-ASP vegetable ASSO time PRT,
(b) *jauh faat-yihn jek-ching-wah.*
then discover CL-frog.
(c) *(wah) jik-haak dong-sih hou ging la.*
(wah) immediately at that time very frightened PRT.
(d) *kam haak-dou chok hai tahng go-douh dou*
so frightened sit from chair there
*dit-jo lohk laih woh.*
fall-ASP down come PRT.

(e) *kam jek-ching-wah ne jihk-dou beih go-neuih-yan*
so CL-frog PRT also PASS CL- woman
*ge jin-giu-sing haak-chan la,*
ASSO scream frightened,
(f) *kam-auk tiu-jo lohk lohk gaak-leih toi jat-go*
so jump-ASP down down next table one-CL
*naahm-jam go-bui-jau jahp-bihn woh.*
man CL-glass-wine inside PRT.
(g) *keuih go-jek-ching-wah ne ne waahk-je ne jan-waih*
3-sg that-CL-frog PRT PRT perhaps PRT because
*toi-jo jahp bui-jau jahp-bihn ne,*
jump-ASP in cup-wine inside PRT,
(h) *jauk jam-jeui-jo la.*
then drunk-PRT PRT.
(i) *kam keuih dak-yihn-gaan san go-tauh cheut-leih,*
so 3-sg suddenly stick CL-head out,

(j) *sek go-go naahm-yan woh.*
kiss that-CL man PRT.

"While the lady was eating some vegetables, (she) discovered the frog. (She) immediately became very frightened. (She) was so frightened that (she) fell from the chair on which (she) was sitting. The frog was also frightened by the woman's screaming. As a result, (it) jumped inside a glass of wine of a man at the next table. Perhaps because that frog jumped inside the glass of wine, (it) was drunk. Suddenly, it stuck (its) head out, and (it) kissed the man."

The frog was mentioned seven times in the above excerpt. The first mention was made in (b). Since prior discourse is not shown, we will not discuss the P.I. and R.D. values of this token.

The second mention of the frog occurred in (e), in which it was re-introduced into the discourse by means of an NP. Since the last mention was made three clauses back, this token of NP was assigned a R.D. value of "3". As for the P.I. level, the token was assigned the P.I. level of "3". The other legitimate alternative for nonhuman referents - the P.I. level of "1" - is not possible with this token because if a zero anaphor is substituted for it, we would not be able to understand which referent is being talked about on the basis of the information provided by (e) and prior discourse: it could be the waiters, the musicians or the other customers in the restaurant that were startled by the woman's screaming.

The frog re-appeared in (f). This time it was reinstated in a zero anaphor. This token of zero anaphora was assigned a R.D. value of "1" since its last mention occurred in the immediately preceding clause. Moreover, it was assigned the P.I. level

of "1" since the non-overt mention is totally unambiguous. Several reasons contribute to the correct identification of the frog, the most interesting reason being that the frog is the only previously-mentioned referent that is capable of jumping into a glass of wine.

The fourth mention of the frog was made in an NP in (g). This token of NP was assigned a R.D. value of "1" and a P.I. level of "1" for the simple reasons that this mention followed immediately the last one, and that the reference is unambiguous.

The frog was mentioned the fifth and the seventh times in (h) and (j) with zero anaphora. Both these tokens were assigned the minimal P.I. level and a R.D. value of "1". The reasons are obvious: they are unambiguous, and both of their last mentions appeared in the immediately preceding clause.

In (i), the sixth case, the frog was reinstated in a third person pronoun. This token was assigned a R.D. value of "1" and the minimal P.I. level. Recall that pronouns in Cantonese can only be used with human referents, yet this pronominal co-reference is not ambiguous. Three crucial reasons for this unusual usage are as follows: first, the current and the last mentions were in close proximity; second, the frog had dominated the discourse for the last four clauses; and third, there were no human referents in the immediate discourse to compete with the frog. These reasons are not independent of each other; they work together to disambiguate the pronominal co-reference.

## 3.5 Results

### 3.5.1 Frequency Distributions of Anaphoric Options in Cantonese Chinese

The twenty narrative samples consist of over two thousand units/clauses (2177), with each of the two stories contributing about half of the total units/clauses. Over one thousand tokens of anaphora were used to reinstate the two boys, the frog, and the monster. The frequency distributions of these tokens are presented in Table 2 on the next page.

The three groups of referents can be differentiated in at least three ways. The first way concerns their relative frequencies of mention by the narrators. The figures in the rightmost column, which is headed "Total", indicate that the two boys as a group were talked about twice as often as the frog and the monster. The reasons are obvious: there are two referents in the human group, who are sometimes talked about separately, whereas there is only one referent in each of the two nonhuman groups;

moreover, the two boys are the focal characters in most of the pictures and are presented in greater detail than are other referents.

| Type / Gram / Referent | Zero Anaphora | | | Pronouns | | | Lexical NP's | | | TOTAL |
|---|---|---|---|---|---|---|---|---|---|---|
| | Sub | Obj | Other | Sub | Obj | Other | Sub | Obj | Other | |
| BOYS | 350 | 0 | 0 | 154 | 11 | 11 | 73 | 4 | 2 | 607 |
| FROG | 90 | 9 | 0 | 14 | 19 | 1 | 105 | 41 | 6 | 285 |
| MONSTER | 24 | 103 | 0 | 2 | 21 | 0 | 36 | 109 | 5 | 300 |
| SUB-TOTAL | 464 | 112 | 0 | 170 | 51 | 12 | 216 | 154 | 13 | |
| TOTAL | 576 | | | 233 | | | 383 | | | 1192 |

Table 2: Frequency Distributions of Anaphoric Options in Spoken Cantonese.

It is not surprising that the monster had a much lower rate of occurrence than the two boys; as is evident from the story, the monster is almost entirely dependent on its creators for its appearance while the same kind of dependency does not hold for the boys.

As for the frog, it tends to have brief appearances in spite of its status as the protagonist of Story F. The non-central human characters that interact with the frog are presented in quite a detailed fashion in the picture book, especially when compared with the minor human characters in the other story. As a result, these non-central characters as a whole were sometimes elaborated to a greater extent than the protagonist in our samples, and they were typically talked about a lot more than the minor characters in the other story.

The three groups of referents can further be distinguished in two important respects. First, they can be differentiated by the feature "humanness", which puts the frog and the monster in a [-human] group and the boys in a [+human] group. The grouping in terms of the feature "humanness" has interesting consequences in relation to the frequency distributions of each type of anaphora under study. On one hand, the two referents in the nonhuman group show very similar distributions of the three anaphoric types: both of them were most often referred to by NP's and least likely reinstated in pronouns. The differences in corresponding proportions for these two referents were tested by chi square; for example, the proportion of zero anaphora for the frog was contrasted with that for the monster. The results of the chi-square test, which are reported in Table 3 below, indicate that none of the differences is critical.

| Anaphor / Ref. / Token | ZERO | PRONOUN | FULL NP |
|---|---|---|---|
| FROG | 99 (35%) | 34 (12%) | 152 (53%) |
| MONSTER | 127 (42%) | 23 (8%) | 150 (50%) |
| $\chi^2$ | 3.56 | 3.02 | 0.65 |
| SIGNIFICANCE LEVEL | Between .10 and .05 | Between .10 and .05 | Between .50 and .30 |

Table 3: Test of Significance of Difference in the Proportions of Each Anaphoric Option Between the Frog and the Monster.

On the other hand, the two nonhuman referents differ from the human ones in significant ways with respect to the frequency distributions in question. While the nonhuman referents were most likely referred to by NP's, the human ones were least often reinstated in explicit anaphora. The chi-square test was performed on the differences in the corresponding proportions for the boys and the frog, and those for the boys and the monster. The results of the test, which are reported in Table 4, show

that all of the differences are highly significant.

| Anaphor / Ref. / Token | ZERO | PRONOUN | FULL NP |
|---|---|---|---|
| BOYS | 350 (58%) | 176 (29%) | 81 (13%) |
| MONSTER | 127 (42%) | 23 (8%) | 150 (50%) |
| $\chi^2$ | 18.92 | 53.33 | 142.12 |
| SIGNIFICANCE LEVEL | Below .001 | Below .001 | Below .001 |

(a) Boys and Monster

| Anaphor / Ref. / Token | ZERO | PRONOUN | FULL NP |
|---|---|---|---|
| BOYS | 350 (58%) | 176 (29%) | 81 (13%) |
| FROG | 99 (35%) | 34 (12%) | 152 (53%) |
| $\chi^2$ | 40.77 | 31.38 | 160.71 |
| SIGNIFICANCE LEVEL | Below .001 | Below .001 | Below .001 |

(b) Boys and Frog

Table 4: Test of Significance of Difference in the Proportions of Each Anaphoric Type for Human and Nonhuman Referents.

The second respect in which the boys, the frog and the monster can be

differentiated is their correlations with certain grammatical cases. The boys strongly correlate with the subject case since ninety-five per cent of the tokens for the boys were found in the subject position. The frog tends to correlate with the subject case but not as strongly as the boys do: only seventy-three per cent of the tokens for the frog occurred in the subject position. The monster shows a clear tendency to correlate with the object case: seventy-eight per cent of the tokens for this referent occurred in the object position.

The correlations observed above were tested by chi-square, which indicates the following results: (1) the boys have a significantly higher chance of appearing in the subject position than the two nonhuman referents regardless of anaphoric types ($p < .001$); (2) the frog shows a significantly better chance of occurring in the subject position than the monster (NP's and zero anaphora: $p<.001$; pronouns: $.01<p<.02$); (3) it follows from (2) and (3) that the monster has a significantly higher chance of correlating with the object case.

The referents' differential tendencies to correlate with particular grammatical cases are not surprising in Cantonese, in which the grammatical subject tends to coincide with the agent role, and the object tends to correlate with the recipient role. The two boys, being human protagonists, play the agent role almost all the time. The scene that contributes most of the objects for the boys is the last scene, in which a group of searchers find the boys in the forest. In contrast with the boys, the monster very often plays the recipient role since it is almost entirely at the disposal of the boys. The frog is quite often the agent of an action, yet it is typically the recipient in quite a number of situations; for example, every time it is discovered by a customer or the personnel in the restaurant, and in the final scenes in which it is being caught by the waiter, rescued and then taken home by its owner.

The correlation between the agent and the subject and that between the recipient and the object are surprisingly strong in this study. In principle, it is possible for Cantonese speakers to use the passive construction - the *"beih- construction"* - to front the recipient to the subject position. Yet, this did not often happen in our data: out of the total tokens of anaphora, we only found eight tokens occurring in the *"beih-construction"*. One of them was an NP used to reinstate the boys, six were NP's for the frog, and one was a pronoun for co-referring with the boys.

To sum up, three general observations have been made with respect to the frequency distributions of anaphoric options in the Cantonese narrative samples

analysed in the present study. The first observation is that the proportion of total tokens for each group of referents serves as a rough indicator of the ways the referents are presented in the picture books: the relative prominence of the referents in the stories and/or the degree of elaboration. The other two observations are as follows: first, the human and nonhuman groups showed differentiated preferences for the three anaphoric options of concern; second, the three groups of referents demonstrated important differences in their tendencies to correlate with certain grammatical cases. These tendencies mirror the semantic role that each group tends to play in the stories.

We now turn to examine the results of the P.I. and R.D. measurements.

### 3.5.2 Referential Distance and Potential Interference

The results of the two measurements will be reported and discussed under the following headings in the following sections: (1) Human Referents; (2) Nonhuman Referents; (3) Zero Anaphora - Subject, Object; (4) Pronouns - Subject, Object, Possessive; (5) NP's - Subject, Object, Possessive.

#### 3.5.2.1 Human Referents

The R.D. and P.I. results for the human zero anaphora are presented in Table 5 below.

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 309 | 20 | 8 | 4 | 7 | 350 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TOTAL | 2 | 309 | 20 | 8 | 4 | 7 | 350 |

Table 5: Zero Anaphora - Human

Moving from the left to the right of Table 5 we first encounter two tokens in the cell with a R.D. value of "1" and a P.I. level of "1". These two tokens were found in relative clauses. The next cell, which has R.D. and P.I. values both equal to "1", contains the vast majority of the human zero anaphora - eighty-eight percent of the

total tokens. Following this cell is the second largest group that consists of about six per cent of the total tokens. The frequency in each of the cells that appear beyond Referential Distance "3" drops below ten or three per cent of the total tokens. No tokens are recorded in the regions with higher P.I. levels.

The frequency distribution of human-referent pronouns is reported in Table 6.

| R.D. / P.I. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 4 | 80 | 17 | 7 | 3 | 11 | 122 |
| 2 | 0 | 25 | 11 | 6 | 3 | 6 | 51 |
| 3 | 0 | 2 | 1 | 0 | 0 | 0 | 3 |
| TOTAL | 4 | 107 | 29 | 13 | 6 | 16 | 176 |

Table 6: Pronouns - Human

An important difference between the pronouns and the zero anaphora for the human referents concerns the overall frequencies of these two types of anaphora: there are only half as many tokens of pronouns as there are of zero anaphora (176 tokens of pronouns versus 350 tokens of zero anaphora). The same kind of discrepancy in the overall frequencies of pronouns and zero anaphora has been observed in written Mandarin by Li and Thompson (1979), who claim that the zero anaphora are the more pervasive form of the two on the basis of the discrepancy.

Another important difference between the two types of non-explicit anaphora for the boys involves the P.I. level of "2". On this P.I. level, fifty-one tokens of pronouns were recorded whereas no tokens of zero anaphora were reported, which implies that the pronouns and zero anaphora for the boys can be distinguished to a certain extent by the factor of potential interference. Nevertheless, the extent is relatively small since only twenty-nine per cent of the pronoun appears on this level.

The majority of the pronouns for the boys (71%) actually falls on the minimal level, where all the zero anaphora are. These pronouns cannot be distinguished from the zero anaphora by the factor of P.I. since they co-occur with the zero anaphora on a

common P.I. level. Can they be differentiated from the zero anaphora by the effect of referential distance?

To answer the last question, we compared the frequencies and proportions of pronouns and zero anaphora at each of the referential distances ranging from "0" to "11[+]" on the minimal P.I. level. The frequency distributions from Distance "5" to Distance "11[+]" are reported in Table 7 below.

| R.D. / ANAPHORA | 5 | 6 | 7 | 8 | 9 | 10 | 11[+] |
|---|---|---|---|---|---|---|---|
| ZERO | 2 | 1 | 1 | 1 | 1 | 1 | 0 |
| PRONOUN | 5 | 0 | 3 | 0 | 1 | 0 | 2 |

Table 7: Frequency Distribution of Zero Anaphora and Pronouns for Human referents from Distance "5" to "11[+]" on the P.I. level of "1".

In our comparisons, we only found significant differences at one referential distance - "1". When the pronouns and the zero anaphora are compared for the boys with respect to their frequencies, the pronouns have a much lower frequency than the zero anaphora at this distance: there are eighty tokens of pronouns at Distance "1", which is only about one-fourth of the tokens of zero anaphora at the same distance. When they are contrasted in relation to their proportions, the pronouns show a significantly lower proportion of tokens: forty-five per cent of the pronouns as opposed to eighty-eight per cent of the zero anaphora for the boys occur at Distance "1" on the mutual P.I. level ($\chi^2_{(1)}=112$, $p<.001$).

At other referential distances on the P.I. level of "1", the pronouns and zero anaphora for the boys show comparable frequencies and proportions. The frequencies and proportions at the zero distance and distances beyond "3" are generally very small - below nine tokens or four per cent for both types of anaphora. The difference in the proportions at each corresponding distance was tested by chi square. The results of the test show that none of the differences reaches the .05 level of significance, which implies that the pronouns and zero anaphora do not show critical differences with

respect to their chances of occurring at the minimal distance and the distances greater than "1" on the minimal P.I. level.

We have commented before that most of the studies in Givon (1983) depend primarily on average R.D. values to study the effect of R.D. on the anaphoric and referential choice, and that averages could be misleading. Our data on the pronouns and zero anaphora for the human referents constitute another piece of evidence in support of our criticism. If we set the upper limit of R.D. value at "20" - the upper limit set by Givon (1983), and then averaged the R.D. values for the pronouns and zero anaphora, we would obtain an average of "2.1" for the pronouns and an average of "1.25" for the zero anaphora. Notice how close these figures are to the prototypical R.D. values reported in Givon's (1983): they are "1" for zero anaphora and from "1.7" to "2" for independent pronouns. If we claim on the basis of the average R.D. values that the human pronouns in our Cantonese data tend to appear at Distance "2", and that they can be differentiated from the zero anaphora along the dimension of R.D., we are very wrong. Only sixteen per cent of the human pronouns occur at Distance "2", and the majority of them (63%) actually fall below this distance!

A last comment about the pronouns for the boys is that nineteen of the tokens that occurred in the subject position were modified by quantifying elements. Specifically, eighteen of these tokens were followed by the quantifier "two", and one token was modified by the quantifier "one", resulting in the following expressions: ***keuih-deih lenhng-go*** (3-pl two-CL) and ***keuih-deih yat-go*** (3-pl one-CL), which mean "they two" and "one of them" respectively.

In spite of the extra information provided by the quantifying elements, no significant differences can be observed between the group of pronouns with quantifiers and the group without such modifying elements: the proportions of tokens on each P.I. level is highly comparable, for example, sixty-eight per cent versus sixty-nine per cent on the P.I. level of "1"; moreover, the modal classes both appear in the "R.D.=1 and P.I.=1" region, and the proportions of the modes are very close (47% versus 45%). Because of the similarities between the two groups of pronouns, they were reported in Table 5 as an undifferentiated group. The results of the group with quantifiers can be found in Appendix B of the present thesis.

To sum up, we have made four observations regarding the frequency distributions of the zero anaphora and pronouns for the boys. First, the pronouns have a lower overall frequency than the zero anaphora. Second, the P.I. level of "2" is

an exclusive region for the pronouns. Third, the majority of the pronouns co-occurs with the zero anaphora on the P.I. level of "1". Fourth, the pronouns have a significantly lower chance of occurring in the "R.D.=1 and P.I.=1" region. We have also illustrated that relying on averages to study the factor of referential distance could be misleading.

Now we proceed to examine the R.D. and P.I. results for explicit anaphora, which are presented in Table 8.

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 4 | 4 | 0 | 1 | 7 | 16 |
| 2 | 0 | 7 | 2 | 0 | 1 | 6 | 16 |
| 3 | 0 | 4 | 2 | 3 | 5 | 35 | 49 |
| TOTAL | 0 | 15 | 8 | 3 | 7 | 48 | 81 |

Table 8: Full NP's - Human

The most obvious difference between the explicit and non-explicit types of anaphora is that over half of the NP's (60%) appears in the potentially most ambiguous region, which implies that disambiguating referents is a crucial function of NP's.

Another obvious difference between the explicit and non-explicit types of anaphora is that the explicit type does not show the same kind of concentration at one single distance or in one cell as the non-explicit type does. The relatively high frequency in the cell with a P.I. value of "3" and a R.D. value of "5[+]" is due to the effect of adding up all the frequencies beyond Distance "4". The individual frequencies that are collapsed under 5[+] in fact are all lower than seven.

On the P.I. level of "2", twenty per cent of the NP's coincides with twenty-nine per cent of the pronouns. The most crucial question to ask about these tokens of pronouns and NP's is again whether they could be distinguished by the factor of R.D.;

specifically, whether the pronouns dominate the lower referential distances while the NP's occupy mainly the higher distances. This kind of division is not observed in the data. The NP's tend to have smaller frequencies than the pronouns at most distances. The difference in the corresponding frequencies of pronouns and NP's is most substantial at Distance "1", where the pronouns outnumber the NP's by eighteen tokens. Beyond this distance, the differences drop below ten tokens and become increasingly smaller as the distances increase. On the basis of these results, we could only claim that there is some evidence for the pronouns to show some advantage over the NP's at Distance "1", and there is no evidence that the NP's have a bigger chance of appearing at higher distances than the pronouns.

These two classes of anaphora were also contrasted with respect to their proportions of tokens at each distance. The difference in the proportions of NP's and pronouns at each distance was tested by chi-square. The results of the test indicate that none of the differences reaches the .10 level of significance, which implies that the two classes of anaphora have comparable likelihood within their own class of occurring at each referential distance on the intermediate P.I. level.

The NP's co-occur with the pronouns and zero anaphora on the minimal P.I. level. The most noticeable difference between the explicit and non-explicit types of anaphora is that the NP's have a much smaller overall frequency: only sixteen tokens of NP's were reported on that level while three hundred tokens of zero anaphora and over one hundred tokens of pronouns were recorded there.

Another obvious difference is that the frequency of NP's at Distance "1" is significantly lower than the frequencies of pronouns and zero anaphora at the same distance. The frequency of NP's at Distance "1" is only five per cent of that of pronouns and slightly over one per cent of that of zero anaphora. As for proportion, five per cent of the NP's as opposed to forty-five per cent of pronouns and eighty-eight per cent of zero anaphora occur at this distance. The differences in the proportions are extremely high, and they were verified by the chi-square test: $\chi^2_{(1)}=43,225$; $p<.001$.

On the basis of the above observations regarding the "R.D.=1 and P.I.=1" region, and an earlier observation that the pronouns have a significantly lower chance of occurring in that region, we can postulate the following hierarchy:

(25) Hierarchy of Preference for the "P.I.=1 and R.D.=1" Region - Human Referents Only

Zero Anaphora > Pronouns > Full NP's

At other referential distances on the minimal P.I. level, the NP's tend to have smaller frequencies than the pronouns and zero anaphora. The most important implication of the last observation is that the NP's do not demonstrate any advantage over the non-explicit anaphoric forms in occurring in regions with high R.D. values on the mutual P.I. level of "1". When proportions are concerned, the three types of anaphora do not show critical differences in the proportions of tokens at each referential distance, which is indicated by the chi-square test. The results suggest that the three classes of anaphora do not differ with reference to their relative tendencies within their classes to appear at each distance other than Distance "1" on the minimal P.I. level.

The proportion of NP's at Distance "5" seems to be significantly higher than that of the zero anaphora at the same distance. This is simply due to the effect of adding up all the frequencies at a distance higher than "4"; when the accumulating effect is removed, the significant difference also disappears.

To sum up, the explicit and non-explicit anaphora for the human referents in this study differ in four important ways and share two similarities. The differences are: (1) the NP's have a much lower overall frequency than zero anaphora and pronouns; (2) the P.I. level of "3" is practically an exclusive level for the NP's; (3) the NP's are less often used to reinstate the boys than are the pronouns at Distance "1" on the P.I. level of "2"; (4) NP's are less preferred than the non-explicit anaphora in the "R.D.=1 and P.I.=1" region. The common properties shared by the explicit and non-explicit anaphora are as follows: (1) the NP's and the pronouns have comparable proportions of tokens at each referential distance on the P.I. level of "2"; (2) at every distance other than "1" on the minimal P.I. level, NP's do not differ significantly from pronouns and zero anaphora with reference to the proportions of tokens at the distance concerned.

3.5.2.2 Nonhuman Referents

The R.D. and P.I. results for the frog and the monster are reported in Tables 9 and 10.

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 93 | 5 | 0 | 1 | 0 | 99 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TOTAL | 0 | 93 | 5 | 0 | 1 | 0 | 99 |

Table 9: Zero Anaphora - Frog

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 105 | 11 | 3 | 0 | 7 | 126 |
| 3 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| TOTAL | 0 | 106 | 11 | 3 | 0 | 7 | 127 |

Table 10: Zero Anaphora - Monster

In general, the frequency distributions of the zero anaphora for the frog, the monster and the boys are quite similar. The modal classes for all three groups of referents are identical - "R.D.=1 and P.I.=1". Moreover, the three modes all contain over eighty per cent of the zero anaphora: eighty-eight per cent for the human referents, ninety-four per cent for the frog, and eighty-three per cent for the monster.

Interesting enough, the proportion of tokens for the frog in the modal class is significantly higher than the corresponding proportion for the monster, which implies that the zero anaphora for the frog tend to appear in the "R.D.=1 and P.I.=1" region more often than those for the monster: $\chi^2(1)=5.50$; $.01 < p < .02$. In contrast, the difference in the proportions for the boys and the frog, and that for the boys and the monster do not reach the .10 level of significance.

Recall that we have asked before if the speaker would be sensitive to the special status of topical/central referents in the hearer's working memory and would therefore tend to re-introduce them in non-explicit anaphora even when their last mentions are relatively far away (Section 2.3.5). Put the question in a slightly different way: would the central/topical referents be reinstated with non-explicit anaphora in regions with high R.D. values more often than the non-central ones because the central referents are remembered better by the hearer? Our data on zero anaphora do not seem to find evidence for such tendency. Our nonhuman protagonist actually turns out to be reinstated in zero anaphora significantly more often than the non-central monster at one of the lowest referential distances - "1"; moreover, the monster, being the least topical referent among the three groups of referents, does not show significantly lower chances of being reinstated in non-explicit anaphora at great referential distances than the central referents.

The lack of evidence for the speaker's sensitivity to the special status of topical referents in our data could perhaps be attributed to the fact that the monster, though not the central character, is quite important in the story. A better test for the sensitivity in question and subsequent adjustments in anaphoric choice might be to compare the protagonists with much less important characters.

We now turn to examine the P.I. and R.D. results for nonhuman pronouns, which are shown in Tables 11 and 12.

| R.D. / P.I. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 14 | 6 | 0 | 0 | 0 | 22 |
| 2 | 1 | 6 | 5 | 0 | 0 | 0 | 12 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TOTAL | 3 | 20 | 11 | 0 | 0 | 0 | 34 |

Table 11: Pronouns - Frog

| R.D. / P.I. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 11 | 2 | 0 | 2 | 0 | 15 |
| 2 | 0 | 6 | 2 | 0 | 0 | 0 | 8 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TOTAL | 0 | 17 | 4 | 0 | 0 | 0 | 23 |

Table 12: Pronouns - Monster

The pronouns for nonhuman referents are "illegitimate": recall that pronouns in Spoken Cantonese are supposed to be for human referents alone. The "illegitimate" status of these pronouns is revealed in several crucial ways. First, they have a low overall frequency, particularly when compared with the pronouns for the human referents (a total of fifty-seven tokens for the two nonhuman referents as opposed to a total of one hundred and seventy-six tokens for the human ones).

Second, the proportions of pronouns for the two nonhuman referents are significantly lower than the proportion of pronouns for the two boys, which has been illustrated before (Section 3.5.1).

Third, the nonhuman pronouns have an overwhelming tendency to cluster in a low R.D. region: ninety-six per cent of them fall below the R.D. of "3" while only seventy-eight per cent of the human pronouns occurs in this low distance region, and the difference in the proportions is significant: $\chi^2_{(1)}=7.86$; $.001<p<.01$.

The most crucial reason for the nonhuman pronouns to occur in a low R.D. region is that it is less difficult for the listeners to figure out which referent(s) these illegitimate pronouns are co-referring with when their co-references/antecedents are in close proximity than when they are not. The factor of proximity is most apparent when the nonhuman pronouns are in the subject position. The subject pronouns, which occur clause-initially, would provide wrong cues to the listener, indicating to him/her that a human referent is being talked about. The effect of wrong signalling is not as detrimental if the last mention of the nonhuman referent concerned is close to the current one. To illustrate the last point, part of the excerpt in (26) is repeated in (27).

(27) (g) *keuih go-jek-ching-wah ne ne waahk-je ne jan-waih*
3-sg that-CL-frog PRT PRT perhaps PRT because
*toi-jo jahp bui-jau jahp-bihn ne,*
jump-ASP in cup-wine inside PRT,
(h) *jauk jam-jeui-jo la.*
then drunk-PRT PRT.
(i) *kam keuih dak-yihn-gaan san go-tauh cheut-leih,*
so 3-sg suddenly stick CL-head out,

"Perhaps because that frog jumped inside the glass of wine, (it) was drunk. Suddenly, it stuck (its) head out, . . . "

The clause in (i) contains the third person pronoun ***keuih***, which co-refers with "the frog". This pronominal co-reference is unambiguous and a few reasons have been offered to account for this (see p.63). One of the crucial reasons is that the last mention of the same referent occurred in the immediately preceding clause: if (h) and (i) were separated by clauses containing referents that are capable of sticking out the head, the pronominal co-reference would be a lot harder to resolve. The final point can be tested by inserting between (h) and (i) the following clause: "The woman sitting at the same table as the man looked absolutely astounded". In this case, the pronoun in (i), which does not carry gender information, and which is a legitimate anaphor for human referents, would very likely be interpreted as co-referring with the woman rather than with the frog.

The above discussion probably explains why only sixteen out of the total fifty-seven tokens of illegitimate pronouns were found in the subject position. Moreover, it is interesting to note that fourteen of these subject pronouns come from the frog, which behaves like a human in the story.

Nonhuman pronouns in the post-verbal object positions do not run into the same kind of problem as those in the subject position. An important reason is that the earlier part of the clause, which includes at least the subject and the main verb, provides extra information to disambiguate the co-reference. In spite of this, speakers are still reluctant to use pronouns to reinstate nonhuman referents in the object position when the last mentions are over two clauses/units away.

The fourth factor which suggests that the nonhuman pronouns are illegitimate is that there are considerable individual differences in the tendency to use the pronouns for nonhuman referents. Let us consider the frog first: two of the speakers who told the story about the frog contribute over half of the total tokens; on the other extreme, one speaker who told the same story did not use one single pronoun to reinstate the

frog at all.

Of the two speakers who account for half of the pronouns for the frog, one of them produced eleven tokens and the other one gave eight tokens. Their respective chances of using the pronouns for the frog are twenty-two per cent and twenty-three per cent. The proportions were obtained by dividing the number of pronouns for the frog by the total number of anaphora for the same referent produced by each speaker and then multiplying the quotient by a hundred per cent. Comparing these proportions separately with the zero proportion; that is, no use of pronouns for the frog, we found significant differences between each set of proportions: $\chi^2_{(1)}=5.45, 4.64$; $.02<p<.05$.

The pronouns for the monster do not show the same kind of bias toward a small number of speakers: the twenty-three tokens of pronouns for the monster were distributed over six speakers whose chances of using the pronouns for the monster range from five to seventeen per cent. The subject with the highest chance is responsible for six of the total tokens. Contrasting the largest chance with the zero chance, we did not find critical differences.

The higher degree of variability shown by the group of speakers who told the story about the frog probably suggests that this adventurous and playful creature induces some speakers to personify it to a greater extent than other speakers. In contrast, the monster, which is not as mobile as the frog and is not really animate, does not tempt the speakers to treat it as a human referent to the same extent as the frog.

The pronouns for the nonhuman and human referents have some properties in common. Three of the major ones are as follows: first, the majority of the tokens occurs on the P.I. level of "1": sixty-five per cent for both the frog and the monster, and sixty-nine per cent for the boys; second, the modal classes for both groups appear in the "R.D.=1 and P.I.=1" region; third, the proportions of pronouns in this class are highly comparable: forty-one per cent for the frog, forty-eight per cent for the monster, and forty-five per cent for the boys. The result of the chi-square test indicates that the differences in these proportions are worse than the .50 level of significance.

To sum up our discussion about nonhuman pronouns, this type of pronoun exhibits the following characteristics: (1) there is a low count of tokens for this class; (2) almost all the tokens were found at a R.D. of below 3; (3) subjects showed a significant amount of individual difference in the tendency to use pronouns to co-refer with the frog; (4) the nonhuman pronouns share two major properties with the human

pronouns: both types of pronouns have a higher tendency to occur on the minimal P.I. level, and the modal classes of both types appear in the "R.D.=1 and P.I.=1" region.

We now proceed to examine the P.I. and R.D. results for the nonhuman NP's, which are presented in Tables 13 and 14.

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 24 | 19 | 14 | 7 | 23 | 87 |
| 3 | 0 | 9 | 8 | 8 | 8 | 32 | 65 |
| TOTAL | 0 | 33 | 27 | 22 | 15 | 55 | 152 |

Table 13: Full NP's - Frog

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 20 | 8 | 9 | 3 | 5 | 45 |
| 3 | 0 | 34 | 21 | 13 | 13 | 24 | 105 |
| TOTAL | 0 | 54 | 29 | 22 | 16 | 29 | 150 |

Table 14: Full NP's - Monster

Four sets of analyses have been performed on the results for the NP's for nonhuman referents. In the first set of analyses, the NP's are contrasted with the zero anaphora for the nonhuman referents. Pronouns are not included in this set of analyses because they are not legitimate anaphora for nonhuman referents.

A major difference between the NP's and zero anaphora for the nonhuman referents is that the maximal P.I. level is almost exclusively dominated by NP's. Only one token of zero anaphora was recorded on this level. This token of zero anaphora was ambiguous to the listener, who asked for clarification right away.

Another difference is that the NP's have a much flatter distribution than the zero anaphora. We have observed that the modal classes of zero anaphora for both referents contain over eighty per cent of the total zero anaphora for each referent. In contrast, the modal classes of NP's for both referents do not contain more than twenty-three per cent of the total tokens for each referent: the modal class for the frog - "R.D.=1 and P.I.=1" - only consists of sixteen per cent of the total NP's for this referent, and that for the monster - "R.D.=1 and P.I.=3" - is made up of twenty-three per cent of the total NP's for the monster. (It is interesting to note that both modal classes of NP's have a R.D. value of "1".)

A further difference between the NP's and the zero anaphora for the nonhuman referents concerns the "R.D.=1 and P.I.=1" region. The frequencies and proportions of NP's for both nonhuman referents in this region are significantly lower than the corresponding frequencies and proportions of zero anaphora. As far as frequencies are concerned, there are twenty-four tokens of NP's as opposed to ninety-three tokens of zero anaphora for the frog; and there are twenty tokens of NP's versus one hundred and five tokens of zero anaphora for the monster. As for proportions, sixteen per cent of the NP's for the frog and thirteen per cent of the NP's for the monster appear in the region concerned. The corresponding proportion of zero anaphora for the frog is ninety-four per cent and that for the monster is eighty-three per cent. The differences in the proportions of NP's and zero anaphora for both referents are highly significant: $\chi^2_{(1)}=144,134$ ; $p < .001$.

Incorporating the results reported in the preceding paragraph with similar results for the human referents which are expressed as the hierarchy in (25), we derive the more general hierarchy in (26) below, with the parenthesis indicating optional elements.

(26) Hierarchy of Preference for the "P.I.=1 and R.D.=1" Region

Zero Anaphora > (Pronouns) > Full NP's

The fourth difference involves the frog only. The frog has a higher chance of being reinstated with an NP than with a zero anaphor in the region with a R.D. value of "3" and a P.I. level of "1". There are fourteen tokens of NP's as opposed to zero tokens of zero anaphora in this region. The difference in the proportions of these two

types of anaphora is significant: $\chi^2_{(1)}=7.99$; $.001< p <.01$.

In the second set of analyses, the two nonhuman referents are contrasted against each other. A general difference between the two referents is that the frog has a higher proportion of tokens on the first P.I. level than the monster: fifty-seven per cent of the tokens for the frog appear on the minimal P.I. level while only about half as many tokens for the monster (30%) occur on this level.

Specifically, the two referents show significant differences in the following three regions: (1) P.I.=1 and R.D.=2; (2) P.I.=3 and R.D.=1; (3) P.I.=3 and R.D.=2.

In the first region, the respective proportions of tokens for the frog and the monster are thirteen per cent and five per cent. The difference in the two proportions is significant: $\chi^2_{(1)}=3.92$; $.02< p <.05$, which suggests that the NP's for the frog have a higher chance of occurring in this region than those for the monster.

In the "P.I.=3 and R.D.=1" region, the monster has a much higher proportion of tokens than the monster does: twenty-three per cent for the frog versus six per cent for the monster: $\chi^2_{(1)}=15.99$; $p<.001$. As of the "P.I.=3 R.D.=2" region, the proportion of tokens for the monster is again significantly higher than that for the monster: $\chi^2_{(1)}=5.67$; $01< p <.02$.

These significant differences between the two referents can be partially attributed to the "*jeung* -construction" in Cantonese. This construction moves the object to a preverbal position, and does not allow object deletion. Since it cannot take zero objects, the P.I. level of "1" is automatically ruled out as an option for this structure.

Scrutinizing the relevant data reveals that ten of the tokens for the monster in the "R.D.=1 and P.I.=3" cell and seven of those in the "R.D.=2 and P.I.=3" cell are objects in the "*jeung* -construction". More importantly, eight of the ten tokens that occur in the first cell and three of the seven tokens in the second cell are unambiguous, and could have been assigned the P.I. level of "1" if the unmarked "non-*jeung*-construction" had been used instead.

If we remove the unambiguous tokens described above from their cells and entered them into the appropriate cells on the P.I. level of "1", we end up with the following new frequencies for the monster: (1) twenty-eight tokens in the "P.I.=1 and R.D.=1" cell; (2) eleven tokens in the "P.I.=1 and R.D.=2" cell; (3) eighteen tokens in the "P.I.=3 and R.D.=2" cell; (4) twenty-six tokens in the "P.I.=3 and R.D.=1" cell. It is interesting to point out that eighteen of the twenty-six tokens in the "P.I.=3 and

R.D.=1" cell were found in the subject position.

Applying the chi-square test to the differences between the new proportions for the monster and the corresponding proportions for the frog, we found these results: (a) the "P.I.=3 and R.D.=1" region indicates a highly significant difference: $\chi^2_{(1)}=8.51$; $.01< p <.001$; (b) the differences in all the other three regions do not reach the .05 level of significance.

The results in the preceding paragraph suggest that the NP's for the frog and the NP's for the monster differ critically only in the "P.I.=3 and R.D.=1" region if structural restriction is removed. The NP's for the monster have a significantly greater chance of occurring in this region, implying that the monster has a higher tendency than the frog to fall in a region which is ambiguous despite a low referential distance.

The higher tendency described above is very likely related to the factor that the majority of the tokens for the monster found in the "P.I.=3 and R.D.=1" cell were grammatical subjects. Recall that the monster tends to correlate with the object case because of the recipient role that it usually plays in the story. It is very likely that this inanimate character faces competition from the human characters, particularly the boys, when it occurs in the subject position; that is, the human characters, which tend to be the grammatical subjects, have a high chance of interfering with the monster when it plays the subject role.

The third set of analyses is concerned with differences between the distributions of NP's for the nonhuman and human referents. We have observed an important difference before: the nonhuman referents have a significantly higher chance of being reinstated in NP's than the human referent.

If we examine first the protagonists of the two stories - the frog and the boys, we can find that they differ critically with reference to their proportions of tokens in the "P.I.=1 and R.D.=1" and the "P.I.=1 and R.D.=3" regions.

In the "P.I.=1 and R.D.=3" region, there are fourteen tokens of NP's for the frog and no tokens of NP's for the boys. The chi-square test was done to test the significance of the difference in the proportion of NP's for the frog and that for the boys, and the difference was found to be significant: $\chi^2_{(1)}=6.39$; $.01< p < .02$.

As for the "P.I.=1 and R.D.=1" region, twenty-four tokens of NP's for the frog are reported in this highly accessible region while only four tokens of NP's for the boys fall in this region. The difference in the proportions of NP's for the frog and the

boys is significant: $\chi^2(1)=4.9$; $.02 < p < .05$.

The critical differences described above probably have something to do with the highlighting effects of NP's and pronouns in unambiguous and typically nondisruptive environments. Li and Thompson (1979, 1981) argue that pronouns in Mandarin play the role of "highlighting" (1981, p.662) when they appear in environments in which zero anaphora can occur. Hinds (1978) reports that one of the functions of the pronouns in Japanese conversation is to "emphasize" (p.174). Consider the excerpt in (28) from one of our speakers:

(28) (a) *yih-ga ne jauh leuhng-go yahn.*
now PRT have two-CL people.
(b) *keuih-deih ne jauh toi-jyuh leuhng-tiuh hou*
3-pl PRT then carry-ASP two-CL very
*cheuhng-ge muhk.*
long-ASSO wood.

"Now there are two people. They are carrying two very long pieces of wood."

The excerpt is the beginning of the narrative by Speaker 103. The speaker introduced the boys in the general term "people" in (a), and reinstated them with the pronoun in the following clause. The pronoun could be replaced by the zero anaphor since neither potential interference nor syntactic constraints on zero anaphora are involved in (b). Yet, it is unusual for the speakers to use the zero anaphor to co-refer with the boys at this early stage of the story. Among the ten speakers that told the story about the boys, only two of them reinstated the boys in the zero anaphor immediately after introducing these two characters. Seven of them used the pronoun and one of them employed the NP to re-introduce the boys.

A possible account of the low tendency to use the zero anaphor to co-refer with the boys at the earliest part of the story is that most of the speakers did not want the boys to disappear totally from the surface forms after their listeners had heard the explicit reference of the two protagonists only once. By reinstating the boys in the pronoun, these speakers prevented the two referents from vanishing from the surface structures, and therefore achieved an effect of highlighting or emphasizing them. Yet, the highlighting effect of the pronoun must be very mild compared with that of the NP since pronouns in Cantonese are quite attenuated.

The highlighting effect achieved by repeatedly using NP's to refer to the same referent when potential interference is not at issue is illustrated in (29).

(29) (a) *kam ne sau-sin ne jauh yauh leuhng-go naahm-jai ne.*
so PRT first PRT then have two-CL boys PRT.

(b)*(kam keuih ne jauh toi-jyuh) leuhng-go naahm-jai ne*
( so 3 PRT then carry-ASP ) two-CL boys PRT
*jauh toi-jyuh yat er yat-di hou cheuhng-ge muhk woh.*
then carry-ASP one er one-CL very long-ASSO wood PRT.

"First of all, there are two boys. (They are carrying) The two boys are carrying one er some very long pieces of wood."

The above excerpt comes from Speaker 106, who was the only speaker that used the NP to reinstate the boys after their introduction. Notice that the speaker first attempted to use the pronoun to reinstate the boys in (b). But she backtracked when she was already half-way into the second verb phrase, and referred to the two boys by an NP.

One of the reasons that might attribute to the repair of speech is that the pronoun in (b) was not properly used since the plural marker was left out. However, this cannot be the major reason because the speaker did not simply correct the pronoun; instead, she replaced it with the NP. She did that probably because she wanted to "double-emphasize" the protagonists at this initial stage of the story so that the listener paid attention to and subsequently remembered them.

If we accept that one of the functions of NP's and pronouns is to highlight a referent, we can argue that the significant differences observed between the protagonists of the two stories with respect to Referential Distances "1" and "3" on the minimal P.I. level is at least partially due to the fact that the pronoun is not as available to the frog as to the boys; that is, the frog has to rely more on NP's to achieve a highlighting effect than do the two human protagonists.

The point made above can be substantiated by an earlier excerpt repeated partially below. Consider the underlined pronoun in (g):

(30) (e) *kam jek-ching-wah ne jihk-dou beih go-neuih-yan*
so CL-frog PRT also PASS CL- woman
*ge jin-giu-sing haak-chan la,*
ASSO scream frightened,

(f) *kam-auk tiu-jo lohk lohk gaak-leih toi jat-go*
so jump-ASP down down next table one-CL
*naahm-jam go-bui-jau jahp-bihn woh.*
man CL-glass-wine inside PRT.

(g) ***keuih*** *go-jek-ching-wah ne ne waahk-je ne*
3-sg that-CL-frog PRT PRT perhaps PRT
*jan-waih toi-jo jahp bui-jau jahp-bihn ne,*
because jump-ASP in cup-wine inside PRT, . . .

"The frog was also frightened by the woman's screaming. As a result, (it) jumped inside a glass of wine of a man's at the next table. Perhaps because it that frog jumped inside the glass of wine, . . . "

Notice that the speaker first used a pronoun to reinstate the frog in (g). He changed his mind immediately after he had uttered the pronoun, and re-introduced the frog by an NP. Though the pronoun is not supposed to be associated with nonhuman referents, the pronominal co-reference in (g) is unambiguous, particularly given the parallelism in the verb phrases in (f) and (g). Since the unambiguous nature of the pronominal co-reference, the repair of speech was very likely caused by the speaker's feeling uncomfortable about using the human pronoun to highlight the frog at that point of the discourse.

We now proceed to compare the other nonhuman referent - the monster - with the boys. The monster shows a higher frequency of NP's in the "R.D.=1 and P.I.=1" region than the boys. There are twenty tokens of NP's for the monster as opposed to four tokens of NP's for the boys in that region. However, when we examined the difference in the proportions of tokens for the monster and for the boys in that region, we found that the difference is not significant: $\chi^2_{(1)}=3.13$; $.05< p<.10$. The observations about the frequencies and proportions seem to suggest that the trend for the monster to have a larger frequency of NP's than the boys in the region concerned is in the right direction since pronouns are not as available to the monster as to the boys. Yet, the monster does not differ critically from the boys with respect to the proportion of tokens in the region in question.

The P.I. level of "3" is not considered in the third set of analyses at all. The reason is that the type of tokens for the nonhuman group of referents recorded on this level is not quite the same as that for the human group: the P.I. level of "3" for the nonhuman group includes tokens that could be assigned the P.I. level of "2" if there

were a category of "nonhuman pronouns".

In the fourth set of analyses, we make the maximal P.I. levels for the human and nonhuman referents comparable by inventing a pronoun for the nonhuman referents. With this hypothetical pronoun, we could eliminate from the P.I. level of "3" all the tokens of the nonhuman NP's that could be replaced by this pronoun. The new distributions of NP's for the two nonhuman referents are reported separately in Tables 15 and 16.

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 24 | 19 | 14 | 7 | 23 | 87 |
| 2 | 0 | 3 | 2 | 1 | 1 | 1 | 8 |
| 3 | 0 | 9 | 8 | 8 | 10 | 33 | 57 |
| TOTAL | 0 | 36 | 29 | 23 | 18 | 57 | 152 |

Table 15: Full NP's - Frog (With P.I.=2)

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 20 | 8 | 9 | 3 | 5 | 45 |
| 2 | 0 | 16 | 8 | 6 | 4 | 7 | 41 |
| 3 | 0 | 18 | 13 | 7 | 9 | 17 | 64 |
| TOTAL | 0 | 54 | 29 | 22 | 16 | 29 | 150 |

Table 16: Full NP's - Monster (With P.I.=2)

The results in Table 15 indicate that fifty-seven tokens of NP's for the frog fall in the maximally ambiguous environment. This portion of tokens is exactly one-fifth of the total tokens of anaphora for the frog. The results in Table 16 show that the P.I.

level of "3" contains sixty-four tokens of NP's for the monster, which is about twenty-one per cent of the total tokens of anaphora for the monster.

In contrast with the nonhuman referents, the human referents have forty-nine tokens of NP's on the P.I. level of "3", or eight per cent of the total number of anaphora. The proportion of tokens for the boys in the maximally ambiguous environments is significantly lower than those for the frog and the monster. The differences in the proportions were tested by chi square, which indicates that the differences are very significant: boys and frog: $\chi^2(1) = 26.35$; boys and monster: $\chi^2(1) = 32.37$; $p < .001$.

The results concerning the maximally ambiguous environments imply that the nonhuman referents tend to fall in those environments more often than the human ones. These results combined with those concerning the low availability of the pronouns to the nonhuman referents could account to a considerable extent why the nonhuman referents show a significantly higher chance of being reinstated in NP's than the human referents.

A last comment about the results in Tables 15 and 16 is that the frog does not seem to be as affected by the availability of the intermediate P.I. level as the monster does. Only eight tokens for the frog could be replaced by the newly-invented pronoun while forty-one tokens of NP's for the monster could be substituted for by the pronoun.

Among the forty-one tokens for the monster, over half occur in syntactic structures that have constraints over zero arguments: nineteen tokens are objects in the "*Jeung*- construction", three tokens are the mutual elements in the pivotal construction, and two tokens co-occur with coverbs. As for the frog, two of the eight tokens that could be substituted for by the artificial pronoun appear in the "*Jeung*- construction", and two tokens are arguments of coverbs.

To sum up, we have analyzed the distribution of the NP's for the nonhuman referents in four ways. First, we compared the NP's and zero anaphora for the nonhuman characters and the major results are as follows: (1) the P.I level of "3" is practically an exclusive level for the NP's; (2) NP's have a significantly lower chance of appearing in the "R.D.=1 and P.I.=1" region than zero anaphora; (3) the frog is more likely to be re-introduced by an NP at Distance "3" on the minimal P.I. level.

Second, the distribution of the NP's for the frog was contrasted against that for

the monster. Significant differences were found in three regions - "P.I.=1 and R.D.=1", "P.I.=1 and R.D.=2" and "P.I.=3 and R.D.=1", and we have demonstrated that the differences were partially caused by the "*Jeung*- construction", which does not allow zero arguments.

Third, the NP's for the nonhuman referents and those for the boys were compared with respect to the P.I. level of "1". The frog has significantly higher frequency and proportion of NP's than the boys at Referential Distances "1" and "3". The monster shows a higher frequency of tokens than the boys at Distance "1", but the difference in the proportions is not statistically significant. A tentative reason related to the low availability/unavailability of the pronouns for the frog was offered to account for the significant differences between the frog and the boys.

In the last set of analyses, the frequencies of the NP's for both the nonhuman and human groups of referents in maximally ambiguous environments were contrasted. The nonhuman referents were found to have a significantly higher chance of falling in these environments than the human ones.

We have examined the results of the P.I. and R.D. measurements with respect to the human-nonhuman distinction, and observed similarities and differences between the two groups of referents and within the nonhuman group. In the following sections, we will examine the results of each anaphoric option in relation to grammatical roles.

3.5.2.3 Zero Anaphora - Subject and Object

The frequency distributions of subject and object zero anaphora for all referents with respect to the P.I. and R.D. factors are reported in Tables 17 and 18.

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 416 | 25 | 8 | 5 | 7 | 463 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| TOTAL | 2 | 417 | 25 | 8 | 5 | 7 | 464 |

Table 17: Zero Anaphora - Subject (All referents)

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 91 | 11 | 3 | 0 | 7 | 112 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TOTAL | 0 | 91 | 11 | 3 | 0 | 7 | 112 |

Table 18: Zero Anaphora - Object (All Referents)

The subject zero anaphora differ from the object ones critically in one region alone - "R.D.=1 and P.I.=1". Ninety per cent of the subject zero anaphora appear in this region whereas eighty-one per cent of the object zero anaphora occur there. The chi-square test was performed to test the significance of the difference in the two proportions. The result of the test indicates that the significance level was between the .02 and .01 levels.

3.5.2.4 Pronouns - Subject, Object and Possessive

The R.D. and P.I. results for possessive pronouns are reported in Table 19.

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 4 | 5 | 0 | 2 | 0 | 0 | 11 |
| 2 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TOTAL | 4 | 6 | 0 | 2 | 0 | 0 | 12 |

Table 19: Pronouns - Possessive (All referents)

The possessive pronouns, which only have twelve tokens, constitute a small

negligible group. Nine of these twelve tokens cluster at the two lowest distances on the P.I. level of "1".

The results for the pronouns in the subject and object cases can be found in Tables 20 and 21.

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 83 | 19 | 5 | 3 | 10 | 120 |
| 2 | 0 | 27 | 8 | 3 | 3 | 6 | 47 |
| 3 | 0 | 2 | 1 | 0 | 0 | 0 | 3 |
| TOTAL | 0 | 112 | 28 | 8 | 6 | 16 | 170 |

Table 20: Pronouns - Subject (All referents)

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 17 | 6 | 0 | 3 | 0 | 28 |
| 2 | 1 | 9 | 10 | 3 | 0 | 0 | 23 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TOTAL | 3 | 26 | 16 | 3 | 2 | 0 | 51 |

Table 21: Pronouns - Objects (All referents)

The subject and object pronouns do not demonstrate any critical differences. The proportions of subject and object pronouns in each corresponding set of cells were contrasted and tested by chi square. No critical differences were indicated by the statistical test.

3.5.2.5 NP's - Subject, Object, and Possessive

The R.D. and P.I. results for full NP's in subject, object, and possessive

cases are reported in Tables 22, 23 and 24.

| R.D. / P.I. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 21 | 19 | 10 | 6 | 25 | 81 |
| 2 | 0 | 4 | 1 | 0 | 1 | 6 | 12 |
| 3 | 0 | 24 | 11 | 10 | 12 | 66 | 123 |
| TOTAL | 0 | 49 | 31 | 20 | 19 | 97 | 216 |

Table 22: Full NP's - Subject (All Referents)

| R.D. / P.I. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 22 | 11 | 13 | 5 | 10 | 61 |
| 2 | 0 | 3 | 0 | 0 | 0 | 0 | 3 |
| 3 | 0 | 20 | 18 | 14 | 13 | 25 | 90 |
| TOTAL | 0 | 43 | 27 | 26 | 17 | 32 | 154 |

Table 23: Full NP's - Object (All Referents)

| R.D. / P.I. | 0 | 1 | 2 | 3 | 4 | 5[+] | TOTAL |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 5 | 1 | 0 | 0 | 0 | 6 |
| 2 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 3 | 0 | 3 | 2 | 0 | 1 | 0 | 6 |
| TOTAL | 0 | 8 | 4 | 0 | 1 | 0 | 13 |

Table 24: Full NP's - Possessive and Adverbial Phrase (All Referents).

Table 24 contains thirteen tokens of NP's, which is a very small fraction of the total tokens. One of these thirteen tokens was found in an adverbial phrase and the rest of the tokens occurred in the possessive case. Like the possessive pronouns, the possessive NP's tend to appear at low referential distances.

Comparing the NP's in the subject case with those in the object case, we only found one significant difference. The proportion of object NP's in the "P.I.=3 and R.D.=2" region is significantly higher than that of subject NP's in the same region. We recorded twelve per cent of the object NP's and five per cent of the subject NP's in the region concerned. The difference in the proportions was tested by chi square, and found to be significant: $\chi^2(1)=5.41$; $p=.02$.

The proportion of object NP's seems to be significantly higher than that of the subject NP's in the "P.I.=3 and R.D.=5[+]" region. Again, the difference is due to the effect of adding up frequencies over a wide range of distances. When we unpacked the accumulated frequencies into individual frequencies, we found that the proportions of subject and object NP's at distances higher than "4" are not significantly different. The frequency distributions of subject and object NP's at greater distances are shown in Table 25.

| R.D. / Gram | 5 | 6 | 7 | 8 | 9 | 10 | 11[+] |
|---|---|---|---|---|---|---|---|
| SUBJECT | 12 | 13 | 10 | 1 | 3 | 5 | 22 |
| OBJECT | 6 | 3 | 5 | 0 | 0 | 2 | 9 |

Table 25: Frequency Distributions of Subject and Object NP's.

3.5.2.6 Summary

This section sums up the major findings in the first study of the present thesis. Three main observations regarding the effect of P.I. were made in the study: (1) on the P.I. level of "1", the zero anaphora, pronouns and NP's for all three groups of referents were reported; (2) the pronouns and NP's for the human referents co-occurred on the intermediate P.I. level, and their respective proportions of tokens that appeared there are twenty-nine per cent and twenty per cent; (3) the P.I. level of "3" was almost exclusively dominated by NP's, and the proportions of NP's for the

boys, the frog and the monster are sixty per cent, forty-three per cent and seventy per cent respectively.

In the study, when more than one type of anaphora for the same group of referents appeared on the same P.I. level, the following analysis was done to find out if the co-occurring types of anaphora could be differentiated by the effect of R.D.: the anaphoric types concerned were contrasted with reference to their frequencies and proportions at each referential distance; in particular, the difference in the corresponding proportions at each distance was tested by chi square to see if it is significant. Most of the differences that were indicated by the test to be critical are summed up in Table 26.

The anaphora were further analysed in the following way: the per cent frequencies of each anaphoric option for the different groups of referents were contrasted; for example, the proportion of the NP's for the frog was contrasted with that for the boys at each referential distance on the P.I. level of "1", and the difference in the proportions was tested by chi square. Part of Table 26 contains a summary of the major significant differences found in this analysis.

| P.I. | R.D. | Significant Differences |
|---|---|---|
| 1 | 1 | Zero > (Pronoun) > NP ; Subject Zero > Object Zero<br>$Zero_{Frog}$ > $Zero_{Monster}$ ; $NP_{Frog}$ > $NP_{Boys}$ |
| | 3 | $NP_{Frog}$ > $Zero_{Frog}$ ; $NP_{Frog}$ > $NP_{Boys}$ |
| 2 | 1-2 | Nonhuman Pronoun > Human Pronoun |
| 3 | 1 | $NP_{Monster}$ > $NP_{Frog}$ |
| | 2 | Object NP > Subject NP |
| | In Genera | Max. Ambiguous Envir.: Nonhuman NP > Human NP |

Table 26: Summary of Major Significant Differences in Proportions in Study 1. ">" Indicates "Higher Proportion".

## 3.6 Discussion

The chief goal of the present study is to evaluate the accessibility model. The specific question asked is: "To what extent can the factors of potential interference and referential distance, which the model represents, account for anaphoric choice in the on-line production of narratives?"

As far as the factor of P.I. is concerned, the results of the study suggest that the main function of the factor is to predict where a certain type of anaphora cannot occur: the zero anaphora cannot appear on P.I. levels other than the minimal one; the pronouns cannot occur on the maximal level; and there is no level on which NP's cannot occur.

The twenty speakers in the study were extremely sensitive to the restrictions on each anaphoric option described in the preceding paragraph. Out of the one thousand one hundred and ninety-two tokens analysed in the study, only four tokens "trespassed" on a forbidden level: they were three tokens of pronouns for the boys and a zero anaphor for the monster. The consequence of the "trespassing" is obvious: the anaphora failed to (co)-refer.

When the anaphora were analysed, we could be quite certain about whether or not a token is ambiguous mainly because the hearers were allowed to ask questions during the hearing and they were conscious that they would have to reproduce the story after the hearing. Moreover, their recalls served as an extra avenue for us to find out if they actually comprehended the story or not.

Though the factor of P.I. can predict where an anaphoric option can or cannot occur, it can hardly be the only factor that affects anaphoric choice. If P.I. were the only or an extremely crucial factor, one should be able to observe that a vast majority of the tokens of a particular anaphoric type coincided with its corresponding P.I. level; for example, over seventy-five per cent of the pronouns appearing on the P.I. level of "2". However, that was not observed in our study. All categories of anaphora except the zero anaphora for each group of referents concerned failed to exhibit this kind of correspondence.

It is not surprising that the zero anaphora almost showed a hundred per cent correspondence if we consider that this type of anaphora does not have an alternative P.I. level. Of the rest of the categories, the pronouns for the human referents have the worst correspondence: only twenty-nine per cent of the tokens appeared on the P.I.

level of "2" and the rest of the tokens almost all fell on the minimal P.I. level. The NP's for the frog also exhibit a poor correspondence: they only had a forty-five per cent chance of occurring on the P.I. level of "3". The NP's for the other two groups of referents did not do as badly as those for the frog: sixty per cent of the NP's for the boys, and seventy per cent of those for the monster were reported on the P.I. level of "3". In brief, the factor of P.I. has the least effect on the pronouns in our data and is most effective with the zero anaphora.

Can the other factor - R.D. - take care of the data that are not accounted for by the factor of P.I.? The answer to this question would provide crucial information as to how powerful the accessibility model is in predicting the alternations among anaphoric forms. If the predictive power is high, the factor of R.D. should be able to explain most of the data left unaccounted for by the factor of P.I.. Let us illustrate the last point with the zero anaphora, pronouns, and NP's for the human referents that coincide on the minimal P.I. level: an ideal situation for the model would be that the zero anaphora dominate a certain low distance, for example, Distance "1", the majority of the pronouns occupies a slightly greater distance such as Distance "2", and the distances beyond those associated with the non-explicit anaphora are predominantly NP's.

Two categories of anaphora get close to the ideal situation. They are the NP's and zero anaphora for the frog on the P.I. level of "1", where the zero anaphora predominate Distance "1" while the NP's have significantly higher frequencies and proportions at Distance "3". The differential dominance of a particular distance by the NP's and zero anaphora on the minimal P.I. level is in favour of the model; but the distinction between the two categories is not as well-defined as that in the hypothetical situation. The two categories do not differ significantly in relation to Distance "2" and the distances beyond "3".

At distances greater than "3", the NP's and zero anaphora for the frog behave in the following ways: when the tokens are added up across distances, the NP's show a significantly higher accumulated frequency and proportion; when each distance is considered separately, the two categories of anaphora do not differ critically in the proprtions of tokens they have at each distance, the main reason being that the individual frequencies of the NP's and zero anaphora at each of the distances concerned are generally very small. These observations probably suggest a mild trend for the NP's to have a slightly higher frequencies at distances beyond "3". The trend is

in a direction favourable to the accessibility model.

The other categories of anaphora all depart from the ideal situation in an important and highly consistent way: they do not demonstrate a correlation between greater referential distances and less attenuated forms of anaphora on mutual P.I. levels. Take for example the zero anaphora, pronouns and NP's for the boys that co-occur on the minimal P.I. level. At Distance "1" the zero anaphora demonstrate a significantly higher proportion and frequency of tokens than the other two classes of anaphora, which is in favour of the model. However, at wider referential distances (R.D.>1), the less attenuated categories do not show any significant advantages over the more attenuated ones. Similar observations have been reported of the pronouns and NP's for the boys on the P.I. level of "2", and of the zero anaphora and NP's for the monster on the P.I. level of "1".

On the basis of the preceding discussions regarding the effects of P.I. and R.D. on anaphoric choice the following conclusion can be drawn: the two effects combined are weakest in accounting for the alternations among anaphoric forms that occur on a mutual P.I. level and in regions with R.D. values higher than "1". The accessibility model predicts a positive correlation between the explicitness of an anaphor and referential distance. While support is shown in the present study for the correlation between the more attenuated anaphoric forms with Distance "1", the study does not find strong evidence for the correlation between higher R.D. values and less attenuated anaphora when the effect of P.I. is controlled for. Of the three groups of referents, only the nonhuman protagonist demonstrates some support for the correlation while the other two groups of referents practically provide no evidence at all for such a correlation. The conclusion undermines the predictory/explanatory power of the accessibility model, particularly the effect of R.D. on anaphoric choice.

The effect of R.D. on anaphoric choice demonstrated in the present study is surprisingly weak. Previous studies (Clancy, 1980; Givon1983a, 1983b, 1984; Hinds, 1983) did not find the effect as weak, especially in the alternation between explicit and non-explicit types of anaphora. What possibly causes the disparity between the present and previous studies regarding the effect of R.D. on anaphoric choice? One candidate would be that the anaphoric system in Cantonese is very different from the anaphoric systems in the languages investigated in the earlier studies. Another candidate is that the methodologies used in the present and previous studies are not identical. We would demonstrate below that it is the second factor that

contributes the most to the disparity concerned.

We have commented before that most of the studies reported in Givon (1983) rely primarily on average R.D. values to study the alternations among different anaphoric/referential forms. Examples from this and other studies have been quoted to illustrate that this kind of practice could be misleading. The example used in the present study involves the pronouns and zero anaphora for the human referents, which have an average R.D. value of "2.1" and "1.25" respectively. These values are extremely close to the prototypical values reported in Givon (1983), which implies that we would have come up with similar conclusions about the effect of R.D. on the alternation between zero anaphora and pronouns for human referents if we had relied on averages alone.

Furthermore, we have criticised the fact that all of the earlier studies do not cross-reference the R.D. and P.I. results. An important implication of this practice is that the effect of R.D. is evaluated without controlling for the P.I. effect. In contrast, the present study controlled for the effect of P.I. in the evaluation of the effect of R.D. by comparing the different classes of anaphora that occur on the same P.I. level. The procedure followed in the present study can help us find out if a pure distance account can predict the alternations among various anaphoric options while the procedure used in previous studies cannot do so because it is confounded by the effect of P.I..

The comment made in the preceding paragraph can be substantiated by our data. Consider the pronouns and zero anaphora for the two boys that appear at Distance "2": twenty-nine tokens of pronouns and twenty tokens of zero anaphora at this distance were recorded; moreover, all tokens of the zero anaphora appeared on the minimal P.I. level while seventeen tokens of the pronouns occurred on the minimal P.I. level, eleven tokens appeared on the intermediate level, and one token fell on the maximal level. In our analysis of the effect of R.D., we only contrasted the tokens of the pronouns with those of the zero anaphora on the mutual P.I. level. In doing so, we found a negligible difference of three tokens in the frequencies of the pronouns and zero anaphora with respect to the distance concerned, and did not find significant difference in their respective proportions. In contrast to our procedure, the earlier studies would consider the total numbers of the pronouns and zero anaphora at Distance "2" when the effect of R.D. is being evaluated. The latter procedure yields a bigger difference in the frequencies of the pronouns and zero anaphora, and a highly significant difference in their respective proportions ($\chi^2_{(1)}=16$; $p<.001$).

The procedure of pooling tokens of the same R.D. values across P.I. levels has some fundamental problems in its underlying assumptions. It necessarily presumes that all the tokens on the higher P.I. levels (P.I.>1) will not change their anaphoric category even if the potentially interfering referents are removed from the discourse environment. Take for example the pronouns for the boys at Distance "2": the pooling procedure assumes *a priori* that all the tokens on the P.I. levels of "2" and "3" would still manifest as pronouns if they occurred in unambiguous environments. This presumption is unwarranted: there is clear evidence that the boys were reinstated in zero anaphora at the distance concerned in the absence of potentially interfering referents, which implies that at least some of the tokens of pronouns on the P.I. levels of "2" and "3" would have surfaced as zero anaphora if the environments had not been ambiguous. Since there is no way to tell before the effect of R.D. is determined how many of those pronouns on the higher P.I. levels would still be pronouns and how many would become zero anaphora if the competing referents were eliminated, pooling them with the tokens on the minimal level is unjustified.

If we re-examine the claim that the explicitness of anaphora is correlated with referential distance, which is quoted in (2) and repeated below, it becomes more evident that the distance effect described in the claim is not a pure distance effect for it does not take into consideration the effect of P.I.:

> (31) The wider the gap between present and preceding appearance of the referent/topic in the discourse register, the harder it would be for the *hearer* to identify it.

It seems that the most crucial function of R.D. in the present study is to establish together with the factor of P.I. the hierarchy of preference among the various types of anaphora for the "R.D.=1 and P.I.=1" region. But even in that region the model faces a different type of problem: almost half of the total tokens of the human pronouns (45%) actually occurred in the region though they come second on the hierarchy. If we only consider the tokens of the human pronouns and zero anaphora on the P.I. level of "1", the proportion of the pronouns at Distance "1" goes up to sixty-six per cent whereas the corresponding proportion of the zero anaphora remains eighty-eight per cent. The overlap of the pronouns with the zero anaphora is very substantial and can hardly be explained away by saying that it is an accidental phenomenon.

What are the pronouns doing in that highly accessible region? Neither P.I. nor R.D. could answer this question. We have in fact followed the suggestions made by Li and Thompson (1979, 1981) and Hinds (1978) and tentatively invoked a third factor - highlighting - to account for the alternations between pronouns and zero anaphora in regions where zero anaphora can appear. The exact nature of the factor of highlighting in Cantonese narration and the role it plays in the alternation between the two non-explicit types of anaphora need to be reseached further in future studies. An important area to study would be the linguistic factors that trigger the speakers to use the pronoun to highlight a referent.

A fourth factor, which is related to the third one, has been observed to have some influence on anaphoric choice in the present study; namely, the availability of the pronouns. It was found that the frog shows a significantly higher chance of being reinstated in NP's than the boys in regions that are unambiguous and that have relatively low R.D. values (R.D.=1 and 3). A tentative explanation for the difference is that the pronouns, which are supposed to be exclusively for human referents, are a lot less available to the frog than to the boys. Possibly because of the imbalanced availability of the pronouns to the two groups of protagonists, the speakers in the present study were more ready to use NP's to highlight the frog than they were to emphasize the boys.

The last comment echoes a similar remark made by Clancy (1980). The researcher observed that her Japanese subjects, who cannot use the pronouns to re-introduce characters in a story, were significantly more likely to use NP's to reinstate a newly-introduced referent than the English subjects. She explained the difference between her two groups of speakers by saying that "It is almost as if ellipsis, which preserves no information about a newly introduced referent, is too radical a form [for the Japanese speakers] to use after presenting the listener with only a single noun phrase to establish a new character in his mind." (p.153)

Could there be a fifth factor involved in anaphoric choice? We have asked before if speakers tend to be economical and provide just the right amount of information for the identification of a referent or if they tend to be conservative and mark the referents with more material than is required. The NP's that co-occur with the non-explicit anaphora on the P.I. levels of "1" and "2" and the pronouns that appear on the P.I. level of "1" might be caused by the speakers' attempt to build redundant information

into the discourse. If that were the case, the behaviour did not seem to be consistent because we observed random alternation between the NP's and the non-explicit types of anaphora at relatively high referential distances. Is the inconsistency caused by individual differences in the tendency to provide redundant information? How could individual differences be measured effectively? By merely looking at the data, they do not seem to suggest that those tokens of NP's and pronouns were contributed by a few individuals. Another challenging question is: "How is the building in of redundant information related to the factor of highlightling, which also involves extra information?" Answers to these questions are not attempted in the present study; we hope that they will be addressed in further studies or a more thorough analysis of the present study.

To sum up this section, the accessibility model with its two major component factors was evaluated in the present study. The study found that the effect of R.D. on anaphoric choice was weaker than the effect of P.I. because the former could account for a smaller amount of data than the latter could. Moreover, the factor of R.D. was not as crucial in the present study as in earlier studies, which was attributed to the differences in the methodologies of the two groups of studies. It was demonstrated that the procedures followed widely in previous studies have serious flaws in them: the practice of relying on averages to study alternations among anaphoric forms could be unreliable, and the procedure of pooling tokens across P.I. levels in the evaluation of the effect of R.D. does not measure pure distance effects; the results of the measurement are very likely confounded by the effect of P.I..

The model is weakest in accounting for the following two sets of data: (1) the alternations between the less and the more attenuated forms of anaphora that occur on mutual P.I. levels and regions with R.D. values higher than "1"; (2) the substantial overlap of the pronouns with the zero anaphora in the "R.D. and P.I.=1" region, which is supposed to be the most privileged domain for the zero anaphora.

Two factors other than P.I. and R.D. were invoked to explain some of the data left unaccounted for by the accessibility model. They are highlighting and the availability of the pronouns, whose roles in anaphoric alternations need to be more thoroughly researched in future studies. A fifth factor, the building in of redundant information, was suggested and is hoped to be pursued in later studies.

To conclude, the model cannot cope with the full range of data. It is weak in predicting when the human pronoun would occur, and factors other than the two

proposed by the model are responsible for the alternation between the pronouns and zero anaphora. Moreover, the claim about the correlation between the explicitness of an anaphor and the referential distance is inadequate, especially when the R.D. value exceeds "1".

A last comment about the present study is that the study is limited in several ways. Firstly, the narrative samples were collected in experimental settings in which the speakers were aware that their narrations would be recorded and studied, which might have some unnatural effects on the speaker. Secondly, only three groups of the characters in the two stories were studied, and all three groups were relatively important characters in the stories. Thirdly, the study only involved two stories. It is not unjustified to suspect that the structures of individual stories very likely have some effects on the alternation of anaphoric forms; for example, the complexity of the stories in terms of the number and kind of characters that are involved and how they interact with one another. To obtain a better understanding of the on-line production of anaphora in narration, we need to accumulate more speech samples based on a wide range of stories, and to compare characters with different degrees of significance. Moreover, the most ideal speech samples would be those that are collected in totally natural conditions.

# CHAPTER FOUR

# THE STRUCTURAL MODEL

## 4.1 Introduction

The structural model considers the alternation of anaphoric forms to be a function of the structure of discourse. Discourse is assumed to be hierarchically organized into levels mediated by a central theme. (Chafe 1979, 1980; Grimes, 1978; Hinds, 1979; Kinstch and van Dijk, 1978; van Dijk 1982; *inter alia*). The exact number and terminologies of these levels may vary from one researcher to another, yet the following four levels are generally accepted for narrative discourse: story, episode/paragraph, event/sentence, and idea unit/clause. Each of these levels consists of a theme/topic, and the higher level themes/topics subsume the lower level ones.

The alternation of anaphoric forms is related to the structure of discourse in two general ways. First, it is associated with the thematic status of a character: thematically important characters tend to be coded by non-explicit anaphora whereas thematically unimportant ones tend to be marked by NP's (van Dijk and Kintsch, 1983; Reichman, 1978, 1981). Second, a specific anaphoric form correlates with a particular level. Marslen-Wilson et al. (1982) report that pronouns in English narratives mark the beginning of a new event. Van Dijk (1982) suggests the commencement of a new episode is signalled by nominal expressions; for example, old characters are introduced into a new episode "in full noun phrases instead of pronouns"$_{(p.181)}$. Tomlin (1984-5) makes similar claims as van Dijk. He demonstrated experimentally that NP's are used to introduce an old character into an episode no matter how recently the last mention of the character was made. Moreover, he showed that pronouns are used to maintain a character after it has been introduced into an episode. He links these linguistic phenomena directly to attentional factors.

Tomlin's study is extremely interesting for it relates the alternation of anaphoric forms to cognitive factors. Moreover, some of the claims that it makes are radically different from those made by the accessibility model; for example, the claim that episode boundaries take precedence over recency. The study will be the focus of the rest of this chapter.

## 4.2 Tomlin's (1984-5) Episode-Attention Model

The general hypothesis of the model is repeated in (32) below:

(32) General Hypothesis (Tomlin, 1984-5, p.4)

> "The syntax of reference in discourse production is tied directly to psychological processes of attention as reflected in the episodic organization of natural discourse data."

According to Tomlin, the fundamental claim of the hypothesis is that the alternation of anaphoric options is "a function of attention allocation by the speaker" (p.4); specifically it claims that, "During the on-line process of discourse production, the speaker uses a pronoun to maintain reference as long as attention is sustained on that referent. Whenever attentional focus is disrupted, the speaker reinstates reference with a full noun, no matter how few clauses intervene between subsequent references."(p.4) The term "pronoun" is not clearly defined in the study, but it can be inferred that Tomlin means third person pronouns, zero anaphora and relative pronouns.

The connection between anaphoric choice and processes of attention is assumed to be directly linked to the episodic organization of discourse production. An episode is defined conceptually by Tomlin as "a semantic unit in discourse organization consisting of a set of related propositions governed by a macroproposition or paragraph level theme. It represents sustained attentional effort devoted to the macroproposition and endures until attention is diverted; that is, it is sustained until an episode boundary is reached."(p.6)

Translated in simple language, what Tomlin is claiming is the following: when the speaker's attention is diverted from one macroproposition to another, the configuration of characters associated with the preceding macroproposition is replaced by a new configuration of characters, which may include new and/or old characters. The first mention of an old character after an episode boundary will be made in an NP even if the last mention occurred in the immediately preceding clause. The reason is that there is an episode boundary, which represents an attention shift, intervening between the two adjacent clauses. In contrast, subsequent mentions of the same character will be made in non-explicit anaphora until another boundary is encountered. The speaker does that because his/ her attention is not disrupted within an episode.

In contrast to the accessibility model, Tomlin's model is oriented toward the speaker since the model ties the alternation of anaphoric forms directly to the attention allocation by the speaker. The researcher does not deny that the alternation concerned may have effects on the listener, but his position is that those effects are not

*deliberately* created by the speaker for the sake of aiding the listener to resolve the anaphora; he says: "local syntactic alternations, like the pronoun-noun alternation considered here, may turn out to be linguistic codes which facilitate the cognitive processes required of the listener to build the intended mental representation."(p.5)

The general hypothesis was tested by Tomlin in two experiments. The first test involved thirty English speakers who were presented with a set of slide pictures that they had not seen before. The pictures consisted of five iterations of an identical theme, which is as follows: two nonhuman characters, such as a bird and a cat, encounter each other, and then the larger of the two characters devours the smaller one. (The best way to characterize the picture sequence is that it represents a food chain!) The slides were shown in pairs or one at a time, and the subjects were told to produce a story based on the slides as they appeared. The test took place in the absence of real listeners.

The basic assumption for the test is that "the shutter release cycle of the slide projector represented a sufficiently strong perceptual disruption for the subject that the subject would be forced to re-orient attention in order to continue with the narrative task."(p.10) If the subjects used NP's to reinstate a character after the disruption but pronouns in subsequent mentions, the hypothesis would be confirmed.

In the second test, ten English speakers were shown an 108 second animated videotape which was about the chasing of a fish by a crab. The subjects were asked to produce on-line descriptions of the videotape. Episode boundaries in the tape were supposed to be "triggered by major thematic breaks, operationally defined for this study [Tomlin, 1984-5] as video cuts accompanied by major scenery changes."(p.7) The subjects did the task without listeners and with no prior knowledge of the cartoon.

The results of both tests showed that about eighty-four per cent of the data could be accounted for by the hypothesis. On the basis of these findings, it was concluded that the alternation of NP's and pronouns is tied directly to the attention allocation by the speaker, which is in turn linked to the episodic organization of the narrative discourse.

The findings of the study are exciting; they provide an account of the alternation of anaphoric forms radically different from that offered by the accessibility model. However, there are some problems with the study. The very first one concerns listeners: the subjects did not tell the stories to real listeners. This creates a big problem for the claim that the alternation of anaphoric options is a speaker-oriented

process: one could argue that the subjects behaved in the expected ways because they did not have to deal with a listener. In the least the factor undermines the generality of the study; it cannot conclude that the data indeed reflect what happens in genuine communication contexts, which involve listeners.

There is actually some evidence in the study that the subjects behaved differently if they had a listener in mind. The study reported that a subject imagined she was telling a story to a child while she was performing the required task. It turned out that this single subject was responsible for most of the pronouns that appeared right after episode boundaries; for example, the relative pronouns in (33) (b) :

| (33) | Episode | Clause | Text |
|---|---|---|---|
| | 1 | (a) | Once upon a time there was a butterfly |
| | 2 | (b) | **who** saw a little bird |
| | 2 | (c) | and greeted it. |

The relative pronoun was used to reinstate the butterfly after the second episode boundary. According to the hypothesis, an NP should have been used instead. The relative pronoun should be considered a counter-example to the hypothesis, but this token and other similar tokens by the same subject were not regarded as such in the study, and the justification provided was that this particular subject "regularly read and told stories to children and said that she performed the task as though she were reading a storybook to a child."$_{(p.16)}$ The justification is not convincing: if the alternation of anaphoric forms is indeed cognitively based, there is no reason why the subject's experience as a regular story-teller and the way she performed her task can alter her pychological processes of attention or make her immune to the perceptual disruption triggered by the shutter release cycle.

The absence of listeners makes the testing condition artificial. There is another unnatural element in one of the experiments: the subjects assigned the videotape were required to produce on-line descriptions of the tape, which implies they did the task under unusual time contraints. The question that arises is whether the time constraint coupled with the absence of a real listener would result in linguistic behaviours that are not normally found in real life; for example, more occurrences of ambiguous pronouns, or descriptions that are hard to comprehend. Consider the narrative sample in Appendix C of the present thesis, which is provided in Tomlin's study as a representative protocol. Imagine that it were read to you in 108 seconds, which is the duration of the tape. Could you follow the narration on the first hearing?

The third problem involves the stimulus materials: both the slide and the videotape are about nonhuman referents, which cannot be differentiated by gender in English. Tomlin acknowledged this factor and he eliminated from the data NP's that were triggered by the switching of grammatical subjects because they represented "clear cases of ambiguity resolution" (p.8). The problem is that the "switch-subject" references were not the only case that could be associated with ambiguity resolution in the study. Consider the NP's in (34) (c), which is the beginning of a narrative based on the picture slides:

(34)

| Episode | Clause | Text |
|---|---|---|
| 1 | (a) | There's a butterfly flying |
| 2 | (b) | The bird's talking to the butterfly |
| 3 | (c) | **The bird's** chasing **the butterfly** |

The two NP's in (34)(c) occurred after an episode boundary and were therefore considered instances consistent with the hypothesis. But one could argue that at least one of them was triggered by ambiguity resolution because if they were both replaced by pronouns, it would not be clear which of the characters was doing the chasing and which was being chased.

The excerpt in (34) can be used to illustrate another problem: one could argue that if the characters in the pictures could be distinguished by gender, the subjects might use pronouns instead of NP's after the episode boundary. This problem implies that materials with characters undistinguishable by gender should be avoided in order to provide a cleaner test of the hypothesis.

The final comment about the study is that the researcher does not seem to be consistent: on one hand he argues that the alternation of anaphoric forms is a speaker-oriented process; on the other hand, he considers "switch-subject" references clear cases of ambiguity resolution" and eliminates them from the data. The consideration implies the following two things: first, he agrees that speakers empathize with the need of the listener; second, he suspects that some of his subjects might pretend they were telling the story to an audience. When the suspicion was actually confirmed by one subject, he regards the counter-examples produced by the subject as irrelevant. His inconsistence suggests that he himself is not quite sure whether the syntactic alternation concerned is a speaker or listener-oriented process.

In brief, Tomlin's data might be confounded by four factors: one of the factors

concerns how the data were analysed, and the other factors are all related to the experimental design, which includes the stimulus materials and the procedures of the experiments. Because of the possibility that the data might be confounded, Tomlin's study is not conclusive. A better designed study is needed to test the hypothesis.

Our second and third studies, which replicate Tomlin's study, control for the confounding variables by providing the story-tellers with real listeners, designing the procedure in such a way that the subjects were not put under time constraint, and using stimulus materials that contain referents distinguishable by pronouns. Let us now turn to these two studies.

# CHAPTER FIVE

# STUDIES TWO AND THREE

## 5.1 Study Two

### 5.1.1 Introduction

This is a partial replicate of Tomlin's (1984-5) study on the relationship between episodic organization and the alternative use of explicit and non-explicit forms of anaphora. The latter study defines episodes as "the sustaining of attention on a particular paragraph level theme" (p.4); furthermore, it is postulated that *the differential use of full noun phrases and pronouns* is tied directly to *the attenton allocation by the speaker* during the on-line production of discourse; that is, anaphoric choice is a speaker-oriented process. Two specific claims are made: (1) "During the on-line process of discourse production, the speaker uses a pronoun to maintain reference as long as attention is sustained on that referent." (2)"Whenever attentional focus is disrupted, the speaker reinstates reference with a full noun, no matter how few clauses intervene between subsequent references." (p.4) It is the latter of the two claims that the present study seeks to investigate.

### 5.1.2 Stimulus Materials

There are two sets of materials, one set for the experimental group and the other set for the control group. The materials for the control group are the two picture books used in the first study of the present thesis. Each book is made up of ten sets of pictures arranged in a sequential order with no blank intervening pages. The picture arrangement is schematically represented in the following figure:

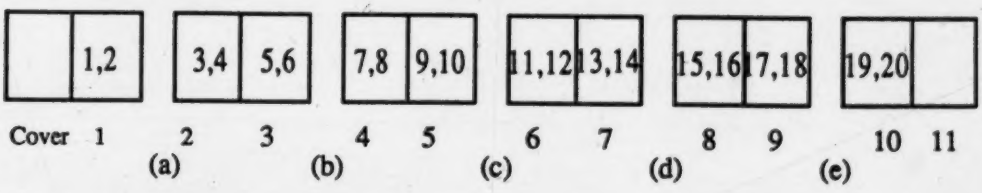



Figure 2: Schematic Representation of Picture Arrangement In Control Condition. Figures In and Beneath Boxes Indicate Picture and Page Number Respectively. Letters Indicate Turning of Pages.

The sequence of pictures is interrupted at Points (a), (b), (c), (d) and (e) when the speaker had to turn over a page to see the following pictures.

The materials for the experimental group are identical to those for the control group in every way except that every pair of pictures is interrupted by a series of five blank pages. The arrangement is schematized in Figure 3.

The purpose of inserting blank pages in the picture sequence is to create artificial episode boundaries that are comparable to those stimulated in Tomlin's (1984-5) study. In that study, episode boundaries are defined operationally as "major disruption in the flow of the nonlinguistic visual material perceived by experimental subjects." (p.7)

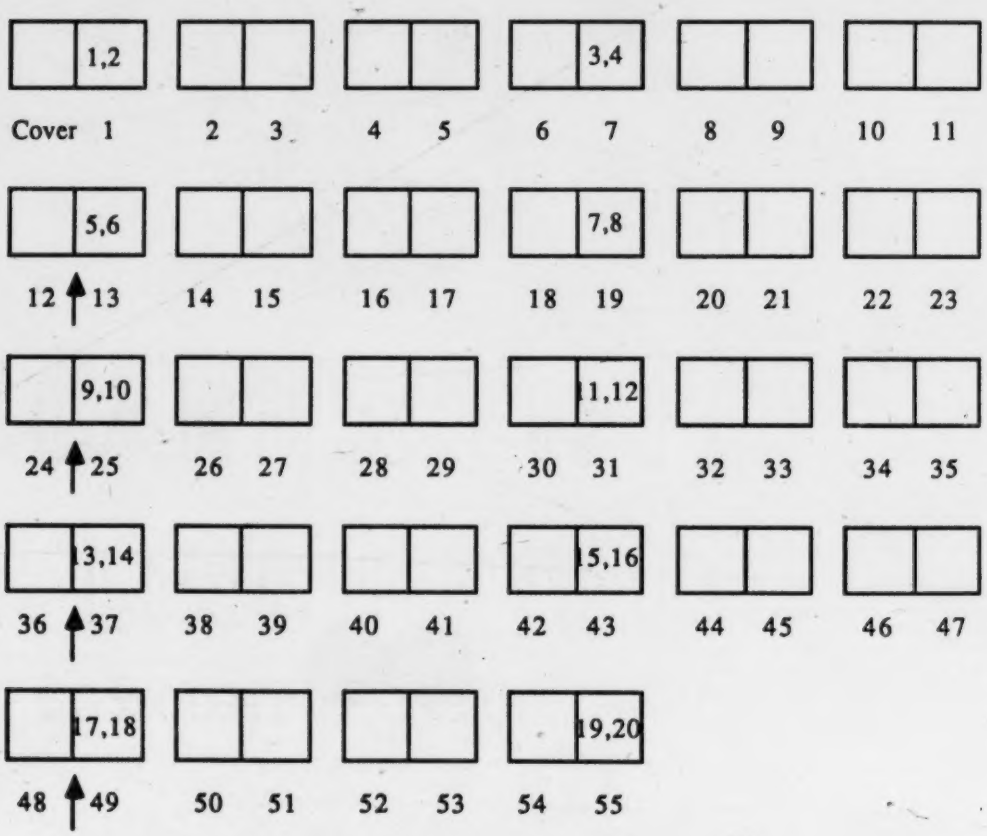



Figure 3: Schematic Representation of Picture Arrangement In Experimental Condition. Figures In and Beneath Box Indicate Picture and Page Number Respectively. Arrows Indicate Testing Points.

Four points in the two picture arrangements which are marked by arrows in the above figure differ significantly with respect to the amount of disruption the speakers received. At these points, the attention of the speakers in the experimental group was disrupted by blank pages while the attention of the speakers in the control group was not disrupted artificially at all because neither blank intervening pages nor turning over of pages was involved. These four points therefore provide the best testing points for the claim that NP's are triggered whenever the attentional flow of the speaker is disrupted.

5.1.3 Subjects

One hundred and twenty Cantonese speakers participated the study. Half of the subjects were assigned randomly to the control condition while the remaining half were assigned to the experimental condition. The control group was the same group of subjects who took part in the first study of the present thesis; the narrative samples collected in that study were used as controls for the present study. Information concerning the control group can be found in Secton 3.2 of the present thesis.

The experimental group was made up of thirty-one male and twenty-nine female natives of Hong Kong, who came from the same high schools as the control group. One third of the subjects played the role of the speaker and the other two-thirds were listeners. Information concerning the subjects' age, sex, and education level is presented in Tables 27 (a-c).

| Age | 17 | 18 | 19 | 20 | 21 |
|---|---|---|---|---|---|
| Speaker | 4 | 7 | 8 | 0 | 1 |
| Listener | 7 | 23 | 7 | 2 | 1 |

(a) Age Distribution

| Sex | Female | Male |
|---|---|---|
| Speaker | 10 | 10 |
| Listener | 19 | 21 |

(b): Sex Distribution

| Education Level | Form VI (Grade 12) | Form VII (Grade 13) |
|---|---|---|
| Speaker | 12 | 8 |
| Listener | 24 | 16 |

(c) Education Level

Tables 27 (a-c): Background Information of Experimental Group in Study II.

Like the control group, the experimental group had received formal instruction in English as a second language. The language was not always used in speaking, but it was used extensively in reading, writing and as a medium of instruction at school. Six of the speakers in the experimental group reported that they could speak one or two Chinese languages or dialects other than Cantonese. Of the six speakers, three of them spoke Mandarin infrequently, one spoke a dialect of the Min language - Fukienese - a lot at home, one knew Mandarin and Fukienese but did not always speak them, and the last one could speak Chiuhjauness - a dialect of the Yue language.

### 5.1.4 Procedures

The procedures in the present study were identical to those followed in Study One with the following exceptions: the experimental subjects were warned ahead of time that there would be blank pages intervening between every two concurrent sets of pictures; moreover, the speakers were told not to skip the blank pages in the process of the telling. If the experimental subjects asked about the purpose of the blank pages BEFORE the test, they were told that the blank pages were there to give the speaker a brief pause every time he/she had finished describing a set of pictures.

Most speakers told the stories to listeners of the same sex. In the first telling, only six of the speakers in the experimental group had a listener of the opposite sex; in the second telling, five of the experimental speakers told the story to a listener of the opposite sex.

### 5.1.5 Transcript

Eighty narrative samples were involved in this study. The twenty samples that were elicited in the second tellings by the control group were transcribed for the first

study of the present thesis, and a detailed description of how the transcripts was prepared can be found in Section 3.4.

The rest of the samples were partially transcribed. The transcribed parts included the first and final three clauses about each pair of pictures. Since the study is only concerned with the portions of narratives occurring immediately after the testing points indicated in Figure 3, we transcribed more than was actually required.

After the transcripts had been prepared, the anaphora that represented the first mentions of the old characters were identified. Those that appeared after the testing points were tabulated under three headings: NP's, pronouns and zero anaphora. NP's that were triggered by ambiguity resolution were eliminated from the data, and all the removed cases involved the switching of the grammatical subjects.

### 5.1.6 Results

#### 5.6.1.1 Overall Results

Two hundred and fifty-two tokens of anaphora were analysed. The distributions of these tokens are presented in Table 28.

| Anaphora / Condition | Full NP | Pronoun/ Zero Anaphor |
|---|---|---|
| Experimental | 38 (.30) | 90 (.70) |
| Control | 14 (.11) | 110 (.89) |

Table 28: Frequency Distributions of Explicit and Non-explicit Types of Anaphora. Figures In Parenthesis Indicate Proportions.

In the experimental condition NP's represent instances that are consistent with the tested claim (hits) while non-explicit anaphora are instances that are inconsistent with the claim (misses). The overall hit rate is only thirty per cent, but it is significantly higher than the proportion of NP's in the control condition: $\chi^2_{(1)}=12.96$; $p<.001$. The results are very interesting: on one hand, they imply that the tested claim is not supported in this study; on the other hand, they suggest that the artificial boundaries had some effect on the alternation of explicit and non-explicit anaphora.

The results will be broken down in the following sections by story type (M or F)

and telling type (first or second).

5.1.6.2 Story M. First Telling

Table 29 contains the proportions of NP's and non-explicit anaphora elicited in the first tellings of the story about the monster:

| Anaphora / Condition | Full NP | Pro/Zero |
|---|---|---|
| Experimental | .22 | .78 |
| Control | .10 | .90 |

Table 29: Proportions of NP's and Non-explicit Anaphora In the First Telling of Story M.

The hit rate is twenty-two per cent and is comparable with the overall hit rate: $\chi^2(1)=1.07$; $.20<p.30$. The proportion of NP's in the control condition seems to be lower than that in the experimental condition, but the difference is not statistically significant: $\chi^2(1)=1.48$; $.20<p<.30$.

The frequency distributions of the tokens by condition and referent type are shown in Table 30.

| Referent / Anaphora / Condition | Central Human | | | Noncentral Human | | | Noncentral Inanimate | | |
|---|---|---|---|---|---|---|---|---|---|
| | NP | Pro | Zero | NP | Pro | Zero | NP | Pro | Zero |
| Experimental | 0 | 24 | 2 | 4 | 3 | 0 | 6 | 0 | 7 |
| Control | 2 | 8 | 12 | 1 | 3 | 6 | 2 | 0 | 15 |

Table 30: Frequency Distributions of Explicit and Non-explicit Anaphoric Options.

Four obserations can be made: first, over half of the hits elicited in the experimental condition come from inanimate referents, which are very hard to reinstate in pronouns, and the rest of them come from noncentral human characters; second, pronouns are strongly preferred for the central human characters; third, the NP's and zero anaphora for the noncentral nonhuman character - the monster - have about an equal chance of occurrence in the experimental condition; fourth, the proportion of zero anaphora in the control condition is significantly higher than that in the experimental condition: $\chi^2_{(1)}=19.31$; $p<.001$, which suggests that more zero anaphora can be used when there is no disruption.

5.1.6.3 Story M, Second Telling

The results for this telling are presented in Tables 31 and 32.

| Anaphora / Condition | Full NP | Pro/Zero |
|---|---|---|
| Experimental | .19 | .81 |
| Control | .02 | .98 |

Table 31: Proportions of NP's and Non-explicit Anaphora.

| Referent / Anaphora / Condition | Central Human | | | Noncentral Human | | | Noncentral Inanimate | | |
|---|---|---|---|---|---|---|---|---|---|
| | NP | Pro | Zero | NP | Pro | Zero | NP | Pro | Zero |
| Experimental | 2 | 16 | 7 | 3 | 1 | 3 | 3 | 0 | 8 |
| Control | 0 | 7 | 13 | 0 | 2 | 8 | 1 | 1 | 12 |

Table 32: Frequency Distributions of NP's, Pronouns and Zero Anaphora.

The following observations can be made: first, the hit rate is low and does not differ critically from the overall hit rate: $\chi^2_{(1)}=1.49$; $.20<p<.30$; second, the proportion of NP's in the control condition is significantly lower than that in the experimental condition: $\chi^2_{(1)}=4.05$; $.02<p<.05$; third, the proportion of zero anaphora in the control condition is significantly higher than that in the experimental condition: $\chi^2_{(1)}=8.02$; $.01<p<.02$.

Recall that the two tellings differ in whether the speakers had prior knowledge fo the story before the tellings. This difference does not seem to affect the control group much since the frequency distributions for the two tellings are quite similar. In contrast, the experimental group shows a statistically significant difference: the subjects used more pronouns to reinstate the central human characters in the first telling than in the second one: $\chi^2_{(1)}=4.48$; $.02<p<.05$. The lower proportion of pronouns in the second telling is probably related to the increase in zero anaphora for the same referents in the second telling: notice that as the number of pronouns decreased from twenty-four in the first telling to sixteen in the second telling, the number of zero anaphora increased from two in the earlier telling to seven in the later telling.

5.1.6.4 Story F. First Telling

The results for this setting are shown in Tables 33 and 34.

| Anaphora / Condition | Full NP | Pro/Zero |
|---|---|---|
| Experimental | .62 | .38 |
| Control | .24 | .76 |

Table 33: Proportions of Explicit and Non-explicit Anaphoric Types.

| Referent / Anaphora / Condition | Central Nonhuman | | | Noncentral Human | | |
|---|---|---|---|---|---|---|
| | NP | Pro | Zero | NP | Pro | Zero |
| Experimental | 10 | 0 | 3 | 0 | 0 | 3 |
| Control | 4 | 5 | 3 | 0 | 0 | 5 |

Table 34: Frequency Distributions of NP's, Pronouns and Zero Anaphora.

The hit rate (62%) in this case is the highest observed so far, yet it does not differ significantly from the proportion of NP's in the control condition (24%) in spite of a thirty-eight per cent difference in the proportions: $\chi^2_{(1)}=3.65$; $.05<p<.10$. The non-significance is due to small frequencies.

Table 34 shows that all the hits come from the central nonhuman character; that is, the frog, which in principle cannot be reinstated in pronouns. The experimental subjects did follow the principle of not using the human pronoun to re-introduce the frog in this telling. In contrast, some subjects in the control group violated the principle: five tokens of the pronoun were used to reinstate the frog. The consequence of the violation seems to be that there are fewer tokens of NP's for the frog in the control condition than in the experimental condition. But we pointed out earlier that the difference in the proportions of NP's between the two conditions is not statistically significant because of small frequencies.

5.1.6.5 Story F, Second Telling

The results for this setting are reported in Tables 35 and 36.

| Anaphora / Condition | Full NP | Pro/Zero |
|---|---|---|
| Experimental | .43 | .57 |
| Control | .27 | .73 |

Table 35: Proportions of Explicit and Non-explicit Types of Anaphora in the First Telling of Story F.

| Referent / Anaphora / Condition | Central Nonhuman | | | Noncentral Human | | |
|---|---|---|---|---|---|---|
| | NP | Pro | Zero | NP | Pro | Zero |
| Experimental | 7 | 1 | 8 | 3 | 1 | 3 |
| Control | 3 | 0 | 10 | 1 | 0 | 1 |

Table 36: Frequency Distributions of NP's, Pronouns and Zero Anaphora.

The hit rate is slightly lower than the miss rate in this case. Most of the hits come from the frog. The proportion of NP's in the control condition is not significantly lower than that in the control condition, which can again be attributed to small frequencies: $\chi^2_{(1)}=0.53$; $.30<p<.50$.

As for the differences between the two tellings, three cells which all concern the nonhuman referent show a difference of at least five tokens: (1) pronoun from the control group; (2) zero anaphora from the control group; (3) zero anaphora from the experimental group. In the last two cases, the frequencies of zero anaphora in the second telling are larger than those in the first telling. But the differences are not statistically significant because the frequencies involved are generally very small.

A last general comment about Story F is that it produces a smaller proportion of

tokens than Story M. The reason is that there are many more "switch subject" cases involved in Story F. Half of the data from Story F were eliminated as a result of that.

In brief, the crucial observations made in this study are as follows. First, the hit rates in three out of the four settings are below chance level. Second, most of the hits from both stories come from nonhuman referents - the frog and the monster, and the noncentral human characters contribute most of the remaining hits. Third, between the two nonhuman characters, the frog has a higher chance of being reinstated in NP's after the disruption. Fourth, the experimental group which told Story M re-introduced the central human characters in pronouns significantly less often in the second telling than in the first one, which is related to an increase in zero anaphora for the same referents in the second telling. Fifth, the control group assigned Story M produced a significantly higher proportion of zero anaphora than the experimental group in both tellings. Sixth, the subjects who told Story F tended to use the zero anaphora to reinstate the frog more often in the second telling than in the first one.

### 5.1.7 Discussion

The results of the study appear to be divided with respect to the tested claim: on one extreme the central human referents show practically no support for the claim; on the other extreme the central nonhuman character seems to provide some evidence for the claim; in between the two ends are the noncentral characters, which show a roughly equal chance of being reinstated in NP's or non-explicit anaphora after the disruption.

The divisions in fact suggest that the claim is inadequate to cope with our data: anaphoric choice after the disruption was not a simple matter of choosing between the NP's and non-explicit anaphora; instead it was based on a general principle interacting with a few factors. The general principle is to avoid using zero anaphora after *considerable* disruption, and the interacting factors involve the availability of pronouns and the information density associated with the characters. The latter factor is determined by the relationship between the characters, the grammatical roles that the characters tend to correlate with, and how firmly the characters had been established in the story prior to the disruption.

Let us consider first the two groups of central characters, which are polarized by Tomlin's claim. These characters share an important similarity: they were rarely reinstated in zero anaphora in the first tellings; in the second tellings more zero

anaphora were used and the reason for the increase will be accounted for later in this section. The two groups of characters differ in one respect: the human group was most often reinstated in pronouns while the nonhuman group was very often re-introduced in NP's. The similarity and difference can be accounted for by the general principle and the availablity of pronouns: the speakers avoided using zero anaphora for both groups of referents after the artificial boundaries and the choice between pronouns and NP's depends on whether the pronouns are available to the particular group of referents; if pronouns are available, the speakers opted for this option; if not, NP's had to be used.

The last point in the preceding paragraph implies that the tendency to use NP's to reinstate the nonhuman character after the disruption was triggered by the low availability of the pronouns to this character. This point is very important because it challenges the support that the nonhuman central character seems to give to the tested claim. It could be argued that the NP's from this character are not real "hits" since the central human characters, which can legitimately be reinstated in pronouns, clearly show that NP's are not the preferred option after the disruption.

We next consider the monster, which is another major source of the "hits" - NP's after the artificial episode boundaries. The NP's are again related to the factor that pronouns are almost not available to the nonhuman characters. In contrast to the central nonhuman character, the monster shows a fifty per cent or better chance of being re-introduced in zero anaphora, which is not predicted by the general principle. This has to do with the following two factors. First, the boys correlate with the subject case whereas the monster correlates with the object case. In fact, all the tokens for the monster recorded in this study were found in the object position. The first factor implies that the boys tended to occur sentence-initially in spoken Cantonese. The second factor is that the monster and the boys always appear together. The two factors put together suggest that the boys, who were always mentioned before the monster, could prime their nonhuman companion. The priming effect probably make the speakers feel less uncomfortable about using zero anaphora to reinstate the monster (McKoon and Ratcliff, 1980).

The observations concerning the nonhuman characters are important because they lend some support to our suspicion that Tomlin's data might be confounded by the factor that all the characters in his stories are nonhuman and cannot be distinguished by pronouns in English. The tendency to reinstate these characters in

NP's after the artificial episode boundaries in his study was very likely triggered by the difficulty in using pronouns to re-introduce them after the disruption.

Our next consideration is the nonhuman noncentral group of referents. These referents almost seem to behave randomly; all three options of anaphora were used for these referents. The randomness is related to a certain degree to the factor that this group consists of the largest number of referents, each of these referents contributing a few tokens. In spite of the irregularities, a generalization could be tentatively made. The choice between NP's and non-explicit anaphora for this group of referents depends on how firmly a particular character had been established in the narrative prior to the disruption: if the character had been quite clearly established, it behaved like the central characters; that is, pronouns rather than NP's were used to re-introduce it after the disruption; otherwise, the speakers marked it with explicit anaphora to make sure that the listeners had no difficulty in identifying it. As for the zero anaphora, they tend to come from the second tellings. This is related to the factor that disruption was alleviated in the second tellings, which will be elaborated later in this section.

Let us return to the general principle and ask why the experimental subjects generally avoided using zero anaphora to reinstate a character after the artificial episode boundaries. The avoidance can be explained from a listener-oriented perspective: the speakers were conscious that the flow of narration had been interrupted by the boundaries considerably, and they helped the listeners to identify the intended referents by providing them with some coding material; moreover, the amount of coding material depends on factors such as the availability of pronouns and the information density associated with a particular character at a specific point in the narrative discourse.

Our account for the general principle can be further supported by two other observations that we have made before. One of the observations concerns the control group which was assigned Story M: this group used zero anaphora significantly more often than the experimental group in both tellings. This must have something to do with the factor that the flow of narration was not disrupted artificially in the control condition.

The other observation involves the second tellings of the two stories: both experimental groups tended to use more tokens of zero anaphora to reinstate the central characters. This is related to the factor that the subjects had seen the stories and knew what was going to happen next, which affects the subjects in two ways: first, some

subjects made anticipatory remarks when they came to an artificial boundary, foreshadowing what the protagonists were going to do in the next set of pictures, and therefore preparing the listeners for the references to these characters after the disruption; second, the time gap between every two sets of pictures was almost reduced by half for both stories in the second tellings. The last point is not surprising if we consider that the subjects had to interpret the pictures in the first telling while they barely had to do so in the second tellings. In brief, this paragraph suggests that the tendency to use more zero anaphora after the artificial boundaries in the second telling is caused by the factor that the disruption was alleviated in this telling. This tendency is most apparent with the central characters and could also be applied to noncentral characters, particularly those that had been firmly established in the narrative prior to the disruption.

To conclude, the claim that speakers use NP's to reinstate old characters when their attention has been disrupted is oversimplistic and can hardly cope with our data. We have offered a more sophisticated account, which has a general principle with a few variables. The general principle says that the speakers avoid using zero anaphora to reinstate a character after considerable disruption unless there is a strong priming effect associated with the character. The choice between pronouns and NP's after the disruption depends on whether the pronouns are available and if the character has been firmly established prior to the disruption.

The present study, which simulated episode boundaries by artificially disrupting the speakers' attention, does not support the claim. Our next study, which is reported below, uses episode boundaries that are inherent in picture sequences presented to the subjects.

## 5.2 Study Three

### 5.2.1 Introduction

The study again investigates the following claim made by Tomlin (1984-5): "Whenever attentional focus is disrupted, the speaker reinstates reference with a full noun, no matter how few clauses intervene between subsequent references." (p.4) Attention disruption is claimed by the reasercher to be triggered by episode boundaries.

The data for this study come from our first study; that is, narratives produced in the second tellings by the control group. This set of data is selected because they most resemble what we will get in real life: the speakers in the test narrated a story to a listener who had not seen the story before, which created a real sense of communication; moreover, the story-teller had prior knowledge of the story, which is again the usual case in real life.

Episode boundaries were determined by getting objective judgements from a large number of Cantonese speakers; moreover, we set a stringent criterion for a well-defined boundary: only boundaries that attained a seventy-five per cent agreement were considered in the study. The reason for the objective judgements and strict criterion is to avoid confounding our results by unclear boundaries. Episodes in general have not been clearly understood nor well-defined in the literature. Their boundaries are usually identified on the basis of thematic unity, which is expressed by unities of character, time, event and location. Chafe (1979) commented that these variables are not static; they interact dynamically to create stronger or weaker boundaries. To make sure we only dealt with the strong cases which most people perceive as episode boundaries, we enrolled a hundred speakers as our judges. The episode boundary judgement task is reported below.

### 5.2.2 Episode Boundary Judgement

#### 5.2.2.1 Stimulus Materials

The two stories for the first two studies of this thesis were used. The pictures on which the stories were based were presented in the form of cartoon strips. Each picture was reduced to about sixty-five per cent of its original size so that eight of them could fit on a single piece of paper fourteen inches long and eight and a half inches wide. There were three such pages which were stapled together. Half of the last page was occupied by a table and some space for the subjects to enter their judgements and information about their sex, age and education level. Instructions were all written in standard Chinese characters. The pictures were numbered and partially coloured so that the figures in them stood out clearly.

#### 5.2.2.2 Subjects

Among the one hundred judges, forty-eight of them were females and fifty-two were males. Ten of the subjects came from one of the high schools that

participated the first two studies of the present thesis, and the rest of the subjects from four high schools that did not take part in the earlier studies. The information concerning the subjects' age and education level is presented in Tables 37 and 38:

| EDUCATION LEVEL | IV | V | VI | VII |
|---|---|---|---|---|
| SUBJECT | 3 | 5 | 53 | 39 |

Table 37: Age Distribution of Episode Judges

| AGE | 17 | 18 | 19 | 20 | 21 |
|---|---|---|---|---|---|
| SUBJECT | 22 | 47 | 24 | 6 | 1 |

Table 38: Education Level of Episode Judges

5.2.2.3 Procedures

Each subject was given a booklet of pictures. Subjects were then instructed by the experimenter to follow these procedures: first, they should examine the pictures carefully until they had understood what the pictures were about; second, they should divide the pictures into meaningful groups entirely on the basis of their own judgements; third, they should enter the judgements in the table provided on the last page of the booklet; fourth, they should fill out the information regarding their sex, age and education level.

The subjects were told to finish the task within fifteen minutes. Moreover, they were allowed to take apart the three pages in the book if they preferred to do so.

5.2.2.3 Results

The results for Stories M and F are presented in Figures 4 and 5 on the following two pages. Only two episode boundaries in each story achieved the seventy-five per cent criterion: one of them is the beginning of the picture sequences;

the other one lies between the fifth and sixth pictures of Story M and between the second and third pictures in Story F.

5.2.3 Anaphor Search

The episode boundaries that met our criterion were marked in the twenty narrative samples and an anaphor search was performed. The search domains were the portions of narratives after the boundaries. The anaphora that represented the first references to the old characters were identified and classified into three groups: NP's, pronouns, and zero anaphora. NP's that were triggered by potential ambiguity were eliminated from the data and they were all caused by the switching of the grammatical subjects.

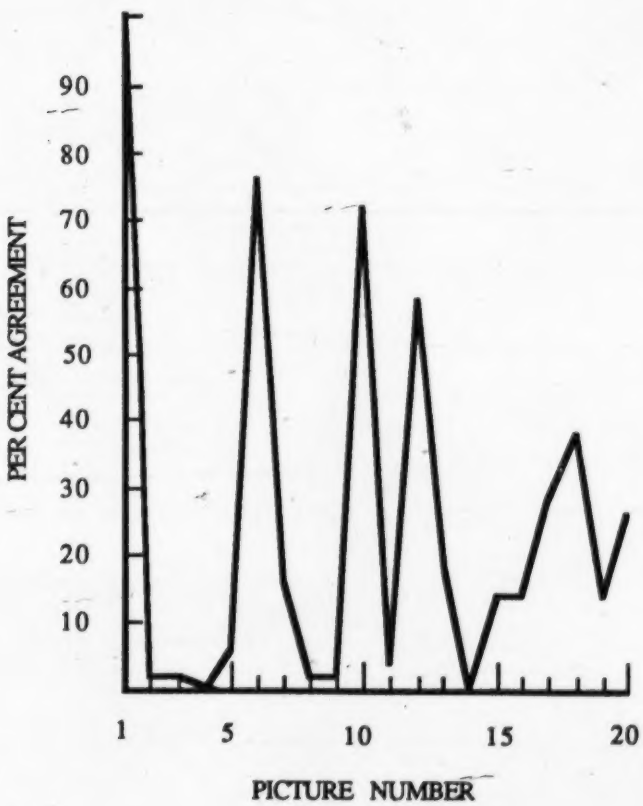



Figure 4: Episode Boundary Judgements on Story M.

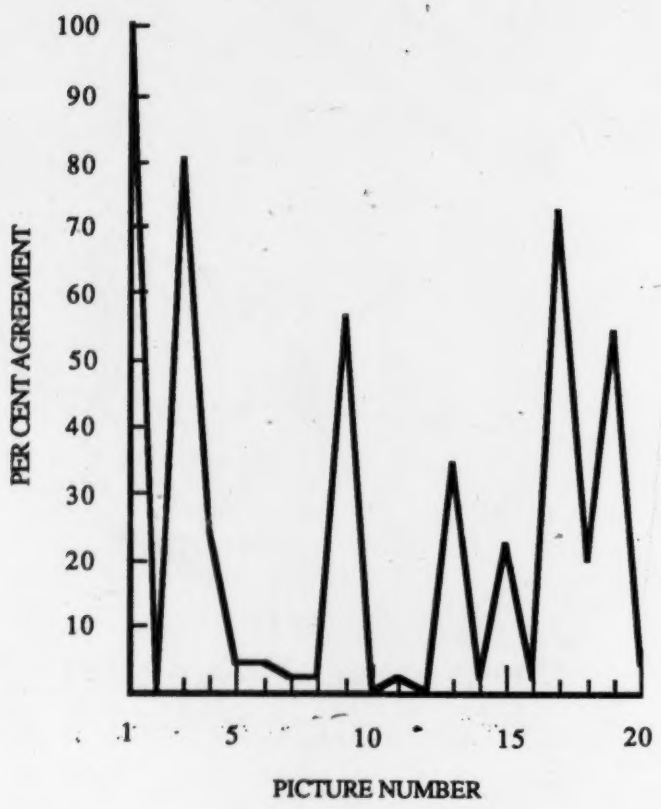



Figure 5: Episode Boundary Judgements on Story F.

5.2.4 Results and Conclusions

The distributions of the anaphora identified in the last section are presented in the following table:

| Boys | | | Monster | | | | Owner of Frog | | |
|---|---|---|---|---|---|---|---|---|---|
| NP | Pro | Zero | NP | | Pro | Zero | NP | Pro | Zero |
| | | | VO | OV | | | | | |
| 0 | 8 | 2 | 1 | 3 | 0 | 5 | 1 | 7 | 1 |

Table 39: Frequency Distribution of Anaphora.

A total of twenty-eight tokens of anaphora was identified. Five of the tokens were NP's, which is about eighteen per cent of the total tokens. Four of the NP's come from the monster and one from the owner of the frog. The single token for the owner was found in repaired speech: the speaker first used a pronoun; then she backtracked and repeated the reference with an NP. Among the four tokens of NP's for the monster, three of them were objects of the "Jeung-construction" (OV), which is a highlighting device in Cantonese, and which does not take zero objects.

The results are exciting: they are extremely consistent with the results of the second study though the data base of the third study is small. First, the general principle is observed: zero anaphora are avoided for both groups of human characters. The owner of the frog, though not the central character in Story F, is the centre of the early part of the story, which includes the initial section of discourse after the episode boundary. This supports our claim that the factor whether or not the noncentral character had firmly been established in the story before the disruption determined the choice between pronouns and NP's. Second, the monster again shows a roughly equal chance of being reinstated in either NP's or zero anaphora. An interesting point about this character is that three out of the four tokens of NP's for this character were associated with a highlighting device in Cantonese, which implies that the choice between NP's and zero anaphora for this character has something to do with whether the speakers wanted to emphasize the monster; that is, when the speakers felt that the monster had an important role to play and should be highlighted, they used NP's to reactivate it after the episode boundaries.

The consistency in the results of the second and third studies suggests that epsiode boundaries are related to the disruption of the flow of discourse because the

boundaries in the second study, which were stimulated artificially, had the same effects on the subjects as the natural boundaries in the third study. Moreover, the consistency suggests that semantic factors such as major changes in themes, and locations have the effect of disrupting the flow of narration: the two natural boundaries examined in the third study both involve transitions from one major theme to another, and are clearly marked by changes in locations and events; changes in these semantic factors had generally caused the speakers in the third study to avoid using zero anaphora to reinstate a character.

To conclude, the third study does not support the claim that NP's correlate with the beginning of an episode though there is evidence for linking natural episode boundaries to the disruption of the flow of narration. Furthermore, the general principle and factors proposed in the second study are valid since they can account for the data of the third study.

Though the claim is not supported in either study, it is important to point out that the structural model has an important implication for us: the discourse structure indeed has an effect on the alternation of anaphoric forms - when the flow of narration is disrupted by artificial or natural episode boundaries, speakers in general avoid using zero anaphora. This suggests that some of the data unaccounted for by the accessibility model might be related to the discourse structure. The data particularly concerned are the human pronouns that appear in the "R.D.=1 and P.I.=1" region, which is a privileged domain for the zero anaphora. The results of our second and third studies suggest that some of these pronouns might be triggered by episode boundaries or other kinds of discourse boundaries; that is, the speakers chose the pronouns when the flow of narration had been disrupted by some kind of boundary.

The possibility that some of the pronouns concerned were caused by discourse boundaries is consistent with a claim made by Li and Thompson (1979, 1981) and Chen (1984): they claim that the personal pronouns in Mandarin have a highlighting effect on a referent, and one of the factors that triggers the need to use pronouns to highlight a referent is the "conjoinability" of two concurrent clauses containing two successive mentions of the same referent: if the conjoinability is low, the pronouns are more likely used; if it is high, zero anaphora are preferred. Some factors that could affect the conjoinability of two clauses are proposed by these researchers; for example, one factor involves temporal adverbials, which are observed to have the effect of impairing the conjoinability of two clauses. We can add to the list of factors a new

member - episode boundaries; that is, conjoinability is greatly impaired when the two clauses are interrupted by an episode boundary.

# CHAPTER SIX

## SUMMARY AND GENERAL DISCUSSION

Two approaches to the alternation of anaphoric forms in the on-line production of narrative discourse have been examined. One approach, called the accessibility model, considers the alternation to be related to the gap between two successive mentions of a referent and the potentiality of interference from other referents. The model was found to be weakest in coping with the substantial overlap of the personal pronouns with the zero anaphora in the "R.D=1 and P.I.=1" region, and the alternation of explicit and non-explicit anaphoric forms in regions with identical P.I. values and R.D. values higher than "1".

The other approach, named the structural model, ties the alternation to the structure of discourse. One aspect of this model has been examined, which is the correlation of NP's with episode boundaries, and which is ruled out as oversimplistic. In spite of that, the structural model is not considered invalid because our data indeed suggest that the alternation of anaphora is influenced by the structure of discourse; specifically, zero anaphora were avoided after considerable disruption of the flow of discourse, which provides a plausible explanation for the overlap of the personal pronouns with the zero anaphora in a region supposedly a privileged domain for the zero anaphora.

The two models turn out to complement each other, the first model capturing some general properties of the syntactic alternation concerned, and the second model pointing to more sophisticated factors that can influence anaphoric choice.

The present thesis has criticized the methodologies used in the studies associated with the two models. The P.I. and R.D. measurements designed by Clancy (1980) and Givon (1983, 1984) are problematic. The R.D. measurements used by both researchers do not measure pure distance effects on anaphoric choice. As for their P.I. measurements, Clancy's measurement defines competing referents in a conservative way and is not justified in its assumption about the correlation between the potentiality of interference and the number of referents intervening between two successive mentions of a referent; Givon's measurement, which only has two levels, cannot cope

important to realize the limitations in Givon's P.I. and R.D. measurements because they have been gaining increasing popularity and have been applied to a growing body of studies: Cooreman (1982), Cooreman, Barbara, and Givon (1984), Verhaar (1984), Rafferty (1985), Rude (1985).

Tomlin's (1984-5) study, which used an experimental approach to study the relationship between episode boundaries and the alternation of explicit and non-explicit forms of anaphora, is criticised for its experimental setting and the stimulus materials; the setting lacks a real sense of communication, and his stimulus materials, which only involve nonhuman referents, are over-restrictive. Notice that if we had used the story about the frog alone, we would have come up with results quite similar to those of Tomlin's and would have missed the general principle that we observed in our second and third studies and the factors that interact with the principle.

However, Tomlin's experimental approach is justified and should be given credit because there is a common problem of circularity in studies that deal with natural text data and attentional factors (see Tomlin, 1983, 1984, 1985 for further comments). For example, Reichman (1978) postulates that NP's are used to refer to referents that are not important to the theme of an episode whereas pronouns correlate with referents crucial to the theme. Her explanation for the correlation is that the speaker's attention is focused on the thematically important referents but not on the thematically unimportant ones. Consider the repeated nominal references to Albert in the following excerpt from Reichman:

(35) (a) one of my problems is that I--I--I put everthing, my feeling, in a total intellectual basis,
(b) I *said* that -- it's funny 'cause, *by the way,*
(c) when I was thinking about *Albert,*
(d) I was thinking how I would think about *Albert,* years from now.
(e) You know look back upon it
(f ) and what context *Albert* would fit in my life.
(g) And my gut phrase was, and *I said, "And I decided that history will really be kind to Albert."*

Reichman's explanation as to why Albert is repeatedly referred to by his proper name is that the speaker's "focus of attention is not on him [Albert] at all" (p.290). The problem with this account is that there is no way the researcher can tell objectively how she knows the speaker's attention is not focused on Albert, apart from giving a circular

answer such as "Because Albert is reinstated in NP's."

Our final comment is that we are aware of the limitations of our three studies with respect to the size of the data, the type of stories involved and the kind of characters analysed. The last two factors have been assessed to be important because we did observe that the characters in the stories show some significant differences with respect to the syntactic alternation concerned. For example, the two nonhuman referents did not respond in the same way to the episode boundaries. Another example is that the two groups of characters in Story M did not demonstrate support for the predicted correlation of NP's with greater referential distances while the frog in Story F showed some evidence in favour of the correlation. These differences suggest that anaphoric choice is probably sensitive to the structure of a story, such as the roles the characters play in the story and the interactions among the characters. More thorough understanding of the linguistic phenomenon concerned is therefore dependent on a larger body of studies using different materials.

## REFERENCES


Anderson, A., Garrod, S. C., & Sanford, A. J. (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology, 35A*, 427-40.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation*. New York: Academic Press, Vol. 8, (pp.47-89).

Bernardo, R. (1976). *The Cognitive Relevance of the Sentence*. Unpublished M.A. thesis, University of California, Berkeley.

Brown, C. (1983). Topic continuity in written English. In T. Givon (Ed.), (pp.313-341).

Caplan, D. (1972). Clause boundaries and recognition latencies for words in sentence. *Perception and Psychophysics, 12*, 73-76.

Chafe, W. (1974). Language and consciousness. *Language, 50*, 111-133.

Chafe, W. (1976). Giveness, contrastiveness, definiteness, subjects, topics, and points of view. In C. N. Li (Ed.), *Subject and Topic*. New York: Academic Press, (PP. 25-55).

Chafe, W. (1979). The flow of thought and the flow of language. In T. Givon (Ed.), (1979). *Syntax and Semantics, vol.12: Discourse and Syntax*.. New York: Academic Press (pp.159-181).

Chafe, W. (Ed.). (1980). *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, N.J.: Ablex.

Chen, P. (1984). *A Discourse Analysis of Third Person Zero Anaphora in Chinese*. Bloomington: Indiana University Linguistics Club.

Clancy, P. (1980). *Referential choice in English and Japanese narrative discourse*. In W. Chafe (Ed.), (1980), (pp. 127-202).

Clark, H. H. & Sangul, C. J. (1979). In search of referents for nouns and pronouns. *Memory and Cognition, 7*, 35-41.

Cooreman, A. (1982). Topicality, ergativity and transitivity in narrative discourse: Evidence from Chamorro. *Studies in Language, 6*, 343-374.

Cooreman, A., Fox, B., & Givon, T. (1984). The discourse definition of ergativity. *Studies in Language, 8*, 1-34.

Daneman, M. & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.

Fletcher, C. R. (1985). The functional role of markedness in topic identification. *TEXT, 5(1-2),* 23-37.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The Psychology of Language.* New York: McGrawHill.

Frederiksen, J. (1981). Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes, 4,* 323-347.

Garrod, S. C., & Sanford, A. J. (1981). *Understanding Written Language: Explorations of Comprehension Beyond the Sentence.* Chichester: John Wiley & Sons. (p.139)

Givon, T. (Ed.). (1983). *Topic Continuity in Discourse: A Quantitative Cross-language Study.* Amsterdam: Benjamin.

Givon, T. (1983a). Topic continuity and word order pragmatics in Ute. In T. Givon (Ed.). (1983), (pp. 141-214).

Givon, T. (1983b). Topic continuity in Spoken English. In T. Givon (Ed.). (1983), (pp. 343-363).

Givon, T. (1984). Universals of discourse structure and second language acquisition. In W. Rutherford (Ed.), *Language Universals and Second Language Acquisition.* Amsterdam: Benjamins, (pp. 109-136).

Grice, H. P. (1967). William Janes Lectures, Harvard Universtiy,1967. Published in part as"Logic and conversation". In P. Cole and J. L. Morgan (Eds.). (1975). *Syntax and Semantics, vol. 3: Speech Acts.* New York: Seminar Press, (pp. 41-58).

Grimes, J. E. (Ed.). (1978). *Papers on Discourse.* Dallas: Summer Institute of Lingusitics Institute.

Halliday, M. A. K. and Hasan, R. (1976). *Cohesion in English.* London: Longman.

Hinds, J. (1978). *Anaphora in Discourse.* Edmonton, Alberta: Linguistic Research.

Hinds, J. (1979). Organizational patterns in discourse. In T. Givon (Ed.), (pp. 135-157).

Hinds, J. (1983). Topic continuity in Japanese. In T. Givon (Ed.), (pp. 43-93).

Hyman, L. (1976). The change from SOV to SVO: Evidence from Niger-Congo. In C. N. Li (Ed.), *Word Order and Word Order Change.* Austin: University of Texas Press, (pp.113-147).

Jarvella, R. (1971). Syntactic processing of connected speech. *Journal of Verbal learning and Verbal Behavior,10,* 409-416.

Jarvella, R. (1979). Immediate memory and discourse processing. In B. Bower (Ed.), (1979). *The Psychology of Learning and Motivation, 13.* New York:

Academic Press, (pp.379-420).

Jarvella, R. J., and Pisoni, D. B. (1970). The relation between syntactic and perceptual units in speech processing. *Journal of the Acoustical Society of America, 48,* 84 (A).

Krahn, F. (1974). *April Fools.* New York: Dutton.

Kintsch, W. and van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363-394.

Li, C. N. and Thompson, S. (1979). Third person pronouns and zero anaphora in Chinese discourse. In T. Givon (Ed.), (pp. 311-335).

Li, C. N., and Thompson, S. (1981). *Mandarin Chinese: A Functional Reference Grammar.* University of California Press.

Maclay, H., and Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word, 15,* 19-44.

Marslen-Wilson, W., Levy, E., and Tyler, L. (1982). Producing interpretable discourse: The establishment and maintenance of reference. In R. J. Jarvella and W. Klein (Eds.), *Speech, Place and Action.* Chichester: Wiley, (pp. 339-378).

Mayer, M. (1974). *Frog Goes to Dinner.* New York: Dial Press.

McKoon, G., and Ratcliff, R. (1980). The comprehension process and memory structures involving in anaphoric reference. *Journal of Verbal Learning and Verbal Behavior, 19,* 668-682.

Murdock, B. B., Jr. (1961). The retention of individual items. *Journal of Experimental Pychology, 62,* 618-625.

Perrin, M. (1978). Who's who in Mambila folk stories. In J. Grimes (Ed.), (pp. 105-118).

Peterson, L. R., and Peterson, M. (1959). Short-term retention of individual items. *Journal of Experimental Psychology, 58,* 193-198.

Piaget, J. (1955). *The Language and Thought of the Child.* New York: World.

Purkiss, E. (1978). *The Effect of Foregrounding on Proniminal Reference,* Unpublished undergraduate thesis, University of Glasgow, Glasgow.

Rafferty, E. (1986). Word order in intransitive clauses in High and Low Malay of the Late 19th Century. To appear in R. N. Tomlin (Ed.), *Grounding and Coherence in Discourse.* Amsterdam: Benjamin.

Reichman, R. (1978). Conversational coherency. *Cognitive Science, 2,* 283-327.

Reichman, R. (1981). *Plain Speaking: A Theory and Grammar of Spontaneous Discourse.* (Tech. Rep.). Bolt, Beranek and Newman, Inc.

Rude, N. (1985). Topicality, transitivity and the direct object in Nez Perce. *International Journal of American Linguistics, 52(2),* 124-153.

Sachs, J. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics, 2,* 437-442.

Sachs, J. (1974). Memory in reading and listening to discourse. *Memory and Cognition, 2,* 95-100.

Sun, C., and Givon, T. (1985). On the so-called SOV word order in Mandarin Chinese: A quantified text study and its implications. *Language, 61 (2),* 329-351.

Tomlin, R. S. (1983). On the interaction of syntactic subject, thematic information, and agent in English. *Journal of Pragmatics, 7,* 411-432.

Tomlin, R. S. (1984). *The Identification of Foreground-background Information in On-line Descriptive Discourse.* Unpublished manucript. University of Oregon, Department of Linguistics, Eugene.

Tomlin, R. S. (1984-5). *The Syntax of Reference in Discourse Production.* (Tech. Rep.). Eugene: University of Oregon.

Tomlin, R. S. (1985). Foreground-background information and the syntax of subordination. *TEXT, 5(1-2),* 85-122.

van Dijk, T. (1982). Episodes as units of discourse analysis. In D. Tannen. (Ed.), *Analyzing Discourse: Text and Talk.* Washington, D.C.: Georgetown University Press, (pp.177-195).

van Dijk, T. & Kintsch, W. (1983). *Strategies of Discourse Comprehension. New York: Academic Press.*

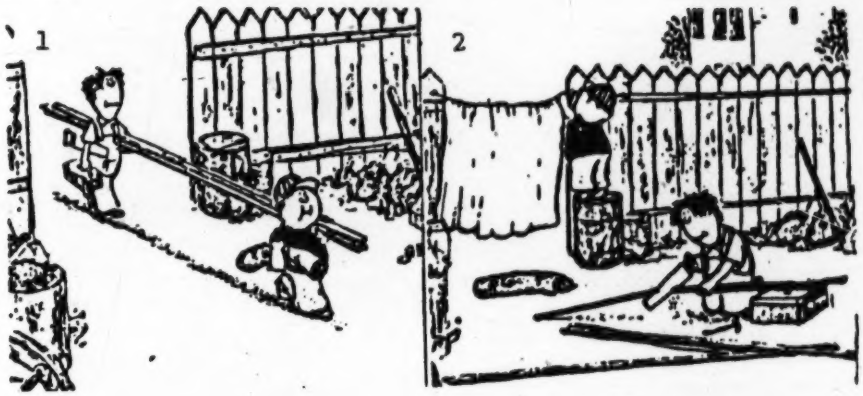
Verhaar, J. Topic continuity in Indonesian. Spokane: St. Michaels Institute, Gonzaga University, MS, 1984.

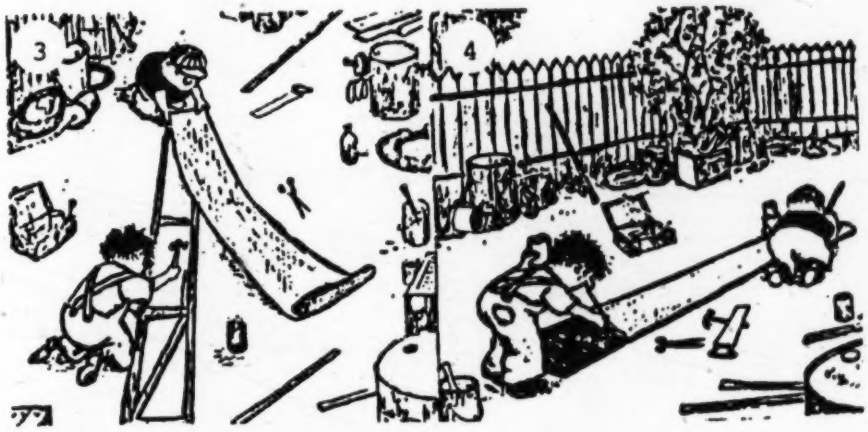
Vygotsky, L.S. (1934). [*Thought and language.*] (E. Hanfman & G. Vaskar, Eds. and Trans.). (1962). Cambridge, Mass.: MIT Press.

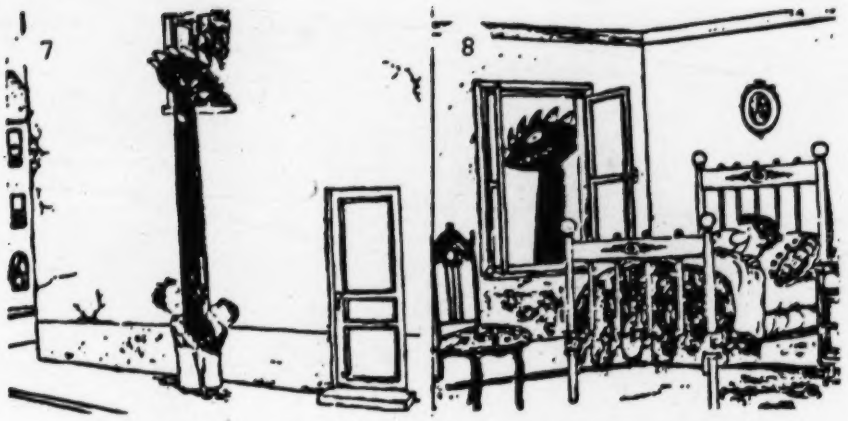
Wieger, L. (1964). *Chinese Characters* (2nd ed., rev.). New York: Paragon Book Reprint Corp.

APPENDIX A

STORY M



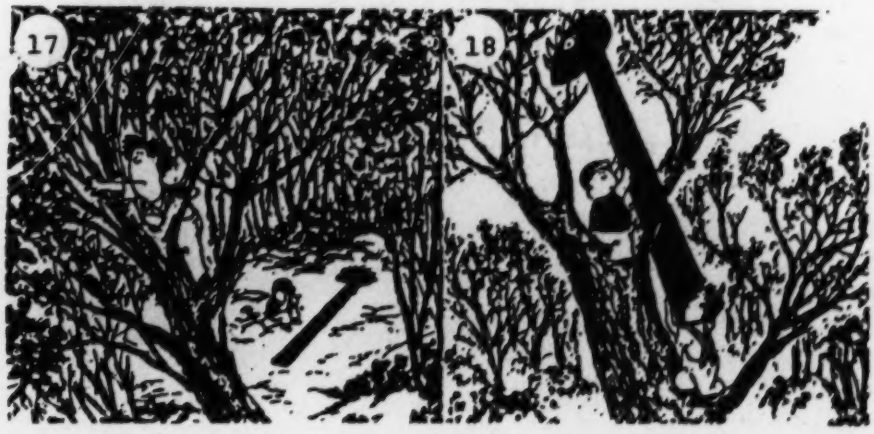

7
8

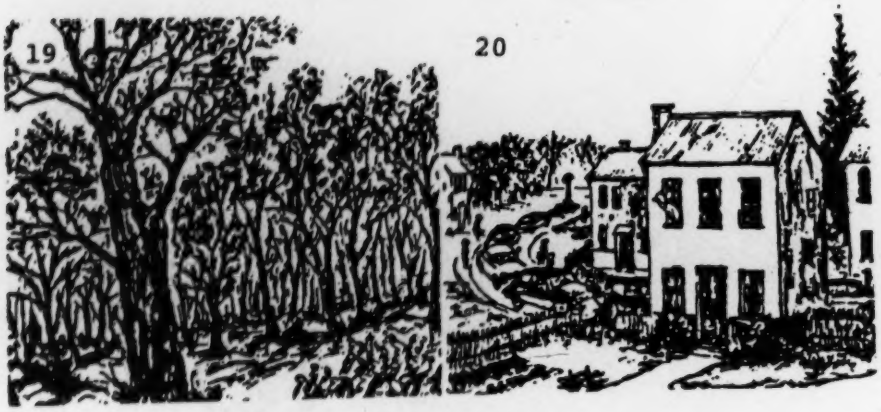
9
10

11
12

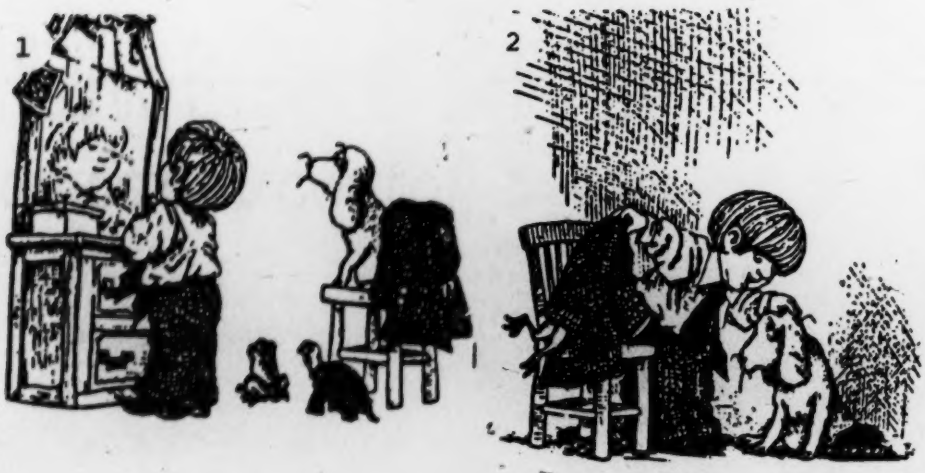
13
14

15
16


17
18




STORY F



3
4

5

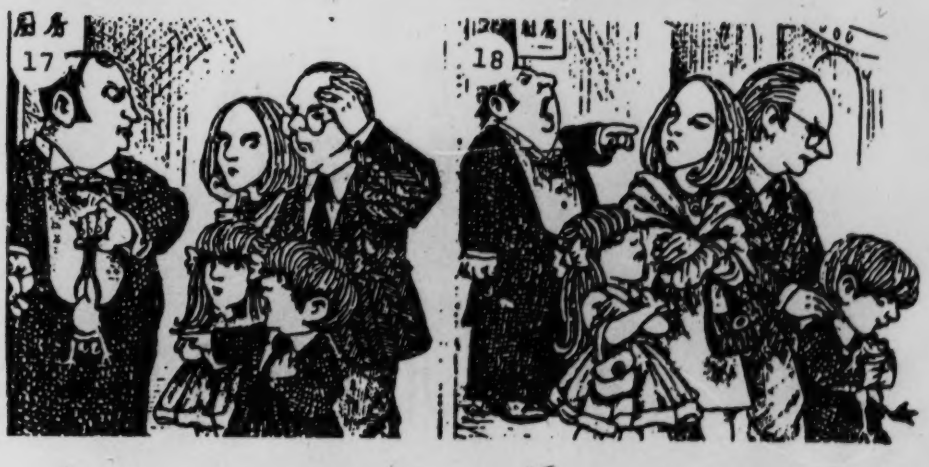
6

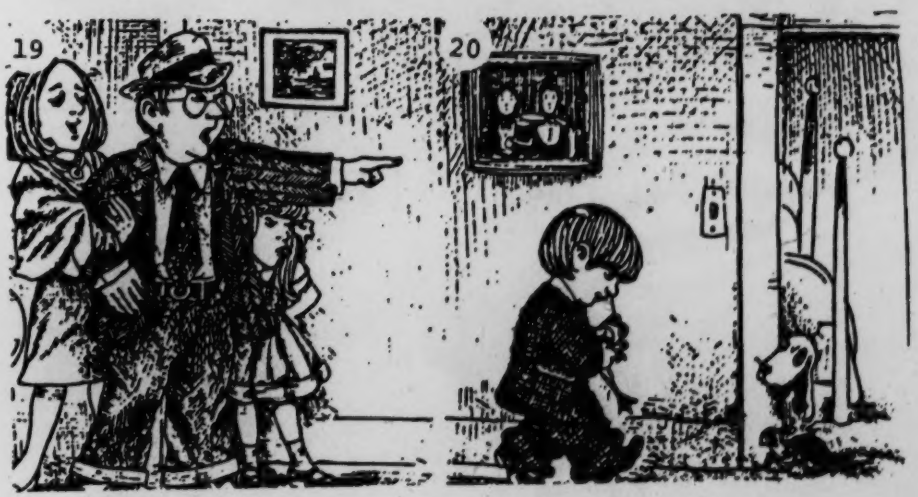
7

8

9
10
11
12
13
14

15
16


17
18


19
20

## APPENDIX B

| P.I. \ R.D. | 0 | 1 | 2 | 3 | 4 | 5[+] |
|---|---|---|---|---|---|---|
| 1 | 0 | 9 | 0 | 1 | 0 | 3 |
| 2 | 0 | 4 | 0 | 1 | 1 | 0 |

Frequency Distribution of Pronouns with Quantifiers for Human Referents.

## APPENDIX C

### A Representative On-Line Narrative Elicited with the Cartoon Videotape from Tomlin (1984-5)

| Episode | Text |
|---|---|
| 1 | Okay, there's a fish in the ocean/ |
| 1 | It's a cartoon/ |
| 1 | And up comes a crab |
| 1 | and tries to get it with its pinchers |
| 1 | and it seems to be avoiding it |
| 1 | and now it disappears/ |
| 2 | Okay, the crab's looking around |
| 2 | and he sees a snail or something |
| 2 | walking on the ocean floor/ |
| 2 | He spots it/ |
| 2 | He blinks his eyes/ |
| 2 | He looks at it |
| 2 | crawling along kind of in a weird fashion, okay/ |
| 2 | Okay, he's been spotted now, the fish that's being chased/ |
| 2 | Okay, he keeps walking/ |
| 2 | Okay, he goes out of the shell |
| 2 | so he's uh lost his uh shell |
| 2 | so the here comes the crab/ |
| 2 | Crawls down from the rock/ |
| 3 | Okay, he lifts his hat up or his shell up to the fish/ |
| 3 | Okay, tries to get him with his pinchers/ |
| 3 | Okay, unsuccessful/ |
| 4 | All right, there are a new set of eyes, a new kind of fish |
| 4 | that seems to be all eyes/ |
| 5 | Okay, the crab's still at it, |
| 5 | trying to catch the fish |
| 5 | and, and then he almost goes in, the fish almost goes in the mouth of a very big whale or shark or something, some some kind of fish/ |

6 Now they're swimming around uh (.) an old wreck of a ship, okay/
6 Obviously been there some time/
6 Um the big fish stretches a little/
6 That was kind of cute/
6 Um, they're still in a chase,
7 the little fish has gone into a bunch of sponges it looks like,
7 and the big fish has gotten lost
7 and crashed into it and/
8 Okay, now he's hidden in a bubble/
8 Okay, so he's, the big fish is going up
8 to try to (.) get in the bubble
8 and catch him
8 and out comes the crab out of the fish's mouth/
8 Okay, they popped the bubble/